Age-Invariant Face Recognition Based on Identity Inference from Appearance Age

Huiling Zhou and Kin-Man Lam

Centre for Signal Processing, Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Kowloon, Hong Kong
E-mail: hl.zhou@connect.polyu.hk, enkmlam@polyu.edu.hk

Abstract

Face recognition across age progression remains as one of the most challenging tasks, as the aging process affects both the shape and texture of a face. One possible solution is to apply a probabilistic model to represent a face simultaneously with its identity variable, which is stable through time, and its aging variable, which changes with time. However, as the aging progress is different for different people, a person may look younger or older than another person, even though their ages are the same. Consequently, using the 'real' age labels given by existing face datasets for age-invariant face recognition will inevitably introduce ambiguity to learning algorithms. In this paper, an identity-inference model, based on age-subspace learning from appearance-age labels, is proposed. We first model human identity and aging variables simultaneously using Probabilistic Linear Discriminant Analysis (PLDA). Then, the aging subspace is learnt independently with the appearance-age labels, and the identity subspace is determined iteratively with the Expectation-Maximization (EM) algorithm. We found that the learned aging subspace is insensitive to the training face images used, and is independent on the identity model. Consequently, the recognition of aging faces becomes simpler as identity inference no longer needs to consider age labels. Furthermore, in our algorithm, different identity features learnt from the identity model are further combined using Canonical Correlation Analysis (CCA), where their correlations are maximized for face recognition. A thorough experimental analysis of face recognition is performed on three public domain face-aging datasets: FGNET, MORPH, and CACD. Experiment results show that the proposed framework can achieve a comparable, or even better, performance against other state-of-the-art methods, especially when the age range is large.

Keywords– Age-invariant, Canonical correlation analysis, Face recognition, Face verification, Identity inference, Probabilistic LDA.

1. Introduction

Face recognition, as a task to identify or verify a person from images and videos, has been intensively studied for decades. With the development of big data and better computational power, general face recognition methods dealing with real-life variations, like poses, expressions, lighting, etc., have achieved superior performances. Latest recognition rates [1] [2] [3] on the most difficult face dataset at present, i.e. Labeled Faces in the Wild dataset (LFW) [4], have been improved to over 99%, which are reported to be even better than human performance. However, face recognition, under age progression, still remains as one of the most challenging problems; the best performance [5], to date, on the most challenging age dataset FGNET [6] stays around 76%. Among the previous related studies on human aging effects [7] [8] [9], it has been widely accepted that facial aging is a complex process, which affects both the shape and texture of a face. In the early growth of a face, from birth to teenager, the greatest change of age progression is in the craniofacial growth (shape change). As people grow older from adulthood to old age, progression of age mainly appears as skin aging (texture change).

Fig. 1. Some face examples from the aging dataset FGNET [6] with real age labels.

There are several reasons why face recognition under age variation is more challenging than other variations: (a). Age progression through life cannot be modeled using a simple progression, as mentioned before; (b). Aging effects are quite specific to

different individuals, so it is almost impossible to precisely define the cause of age progression. For example, healthy people who reach their old age will probably look quite different from those who have suffered from accidents or diseases in their lives; (c). Collecting suitable training data for studying the aging effects is also difficult, as it requires a much longer time period and greater effort. Aging datasets, collected from photos of different age stages, may undergo more serious distortion than other variations, as shown in Fig. 1; and (d). Last but not least, almost all the previous age-related research work is based on datasets where a real-age label is given to each individual. This makes the recognition task by machines incredibly tough, because most of the existing methods can only teach machines to learn from the facial-appearance information. Two people with similar real ages may look very different in appearance, as shown in Fig. 1(b) and 1(d). It will inevitably make the learning or classification process less accurate.

In recent years, many age-related works have been proposed on age estimation [10] [11] [12] [13] [14], age simulation [15] [16] [17] [18], age-invariant face recognition or verification [5] [19] [20] [21] [22], etc. While they serve different application goals, the underlying theories and methods overlap and correlate extensively. Generally, all these approaches can be categorized into two groups. The first is the generative approaches [10] [15] [19] [21], which construct 2D or 3D generative models to compensate for the aging process, and synthesize face images that match the age of query face images. The second approach is based on discriminative models [20] [22] [23] [24] [25], which use robust facial features and discriminative learning methods to reduce the gap between face images captured at different ages.

Age estimation and simulation both use similar approaches to age-invariant face recognition tasks. However, age estimation and simulation mainly focus on manipulating the aging information that varies with age progression, while age-invariant face recognition aims to seek the identity information that is stable for the same individual over age progression. This substantial difference inspires a new approach that attempts to separate a face into its aging factor and identity factor [18] [21] [26]. One of the earliest works on face recognition that describes a face with its within-individual

and between-individual variations was introduced in [26] [27] [28]. Probabilistic Linear Discriminant Analysis (PLDA) [42] was employed to establish a generative linear model, and the optimal latent identity variable was iteratively derived by using the Expectation-Maximization (EM) [29] algorithm. This method was further applied to age-invariant face recognition in [21], where the within-individual variance was suitable for using the aging information, while the between-individual variation was suitable for using the identity information. Again, the EM algorithm is used to obtain both the latent variables simultaneously, and the identity factor is then used for recognition. Experiments showed that this method outperforms other existing methods. Later on, this idea was also applied to render aging faces, by modeling the aging layer as a linear combination of age-progression patterns while keeping the personalized layer invariant through time [18]. All these methods generate the aging subspace and the identity subspace using a single model at the same time. However, this approach has a high demand on the training datasets, because both the identity and the aging information must be learnt as thoroughly as possible. Unfortunately, it is a great challenge to obtain suitable datasets for age-invariant face recognition. For the three most well-known datasets for this task, they either suffer from lack of training samples (FGNET dataset [6]) or lack of samples with long time periods for learning aging patterns (MORPH [30] dataset and CACD dataset [31]). What's worse, all the previous learning frameworks were based on real-age labels, which may be inconsistent with the corresponding appearance ages (people with the same real age may look different in age due to differences in individual skin care or health conditions). This means that the existing methods achieve limited performances on face recognition with age variations. One way to solve the age-gap problem is to seek the underlying sequential patterns [10] [32], and then apply manifold learning to analyze the age characteristics [11] [33]. It has been shown that applying the orthogonal Locality Preserving Projections (OLPP) [34] to an aging database, with ages ranged from 0 to 93 years, yields better statistical age estimation results.

With the increasing interest in age-related topics, the corresponding tasks have become more and more demanding in recent years. The appearance age estimation

challenge on the public ChaLearn dataset [35], with face images in the wild and labeled with the appearance age, is one of the great sources for learning age progression. In this paper, we propose an age-invariant face recognition framework, namely aging-guided identity inference model (AG-IIM), where the feature gap between two face images of the same person, captured at different ages, can be reduced. Similar to the method that deals with the aging and identity information separately [18] [21] [26], we establish a generative model based on PLDA. Unlike those previous works, which learnt and derived the aging and identity subspaces at the same time, we propose to learn aging subspace separately by using manifold learning on an aging dataset with appearance-age labels.

The contributions of this paper are given as follows. First, we empirically show that our method can obtain a more discriminative identity subspace. Second, our method is the first to tackle the age-invariant face recognition problem based on appearance age, so that computers can both learn more effectively and more consistently. What's more, a byproduct of this framework is to provide a much easier way to collect aging photos for face recognition, where only identity labels are required. The aging characteristics can be learnt by any aging dataset with appearance-age labels. Having obtained the identity and aging subspaces, as well as the underlying identity factors based on different features, an effective fusion mechanism based on Canonical Correlation Analysis (CCA) [36] is utilized to further boost the recognition performance. Extensive experiments on three different aging datasets show that our framework can achieve a great improvement in terms of the rank-1 recognition accuracy compared to other state-of-the-art methods, especially when the faces undergo large age variation.

The remainder of the paper is organized as follows. In Section 2, the proposed aging-guided identity-inference model is introduced, where details of the independent aging subspace learning and model optimization will be presented. Section 3 describes the age-invariant face recognition framework, including the training stage, testing stage, and feature fusion for matching. Experimental results and analysis of face recognition with different age labels and on different aging datasets are given in Section 4. The conclusion and future work are outlined in Section 5.

2. Aging-guided Identity-Inference Model

In this section, the proposed identity-inference model, based on PLDA and independent aging subspace learning, is presented. Unlike the previous work that jointly learns the subspaces for aging and identity at the same time, we first derive a discriminative aging subspace by preserving the locality of the appearance-age information, and then the identity subspace with the assistance of the aging subspace through the PLDA model. This strategy makes optimization on the latent variables more efficient, and the collection of aging face images easier, as only the identity label is required.

2.1 Identity-Inference Model

Similar to [21] [26], which use PLDA for face recognition, we also model a face by incorporating both the within-individual variation (aging) and the between-individual variation (identity). Suppose that the nth image of individual m is denoted as x_{mn}, then the identity-inference model can be presented as follows:

$$x_{mn} = \mu + \mathbf{E}u_m + \mathbf{A}v_{mn} + \varepsilon_{mn}, \tag{1}$$

where the first two terms are comprised of the identity components μ and $\mathbf{E}u_m$, which depend only on the identity of the person, while the last two terms are comprised of the aging components $\mathbf{A}v_{mn}$ and ε_{mn}, which are different for images of the same individual and represent the within-individual noise.

Generally, μ represents the overall mean of the training set. The matrix $\mathbf{E}$ is called the identity subspace, whose columns are the bases for cross-identity variations, and u_m can be viewed as the position of x_{mn} in this subspace, which is also the identity factor required for recognition. Similarly, the matrix $\mathbf{A}$ is the aging subspace to depict cross-age variations, and v_{mn} is the corresponding aging factor. The term ε_{mn} represents the remaining noise caused by pose, expression variations, etc., and can be modeled as a Gaussian function with diagonal covariance Σ. The goal of establishing this PLDA model is to compute the likelihood that two face images are generated from

the same underlying identity factor $\boldsymbol{u}_m$ for age-invariant face recognition.

The models in (1) can be re-written in terms of conditional probabilities as follows:

$$Pr(\boldsymbol{u}_m) = G_u[\boldsymbol{0},\mathbf{I}], \tag{2}$$

$$Pr(\boldsymbol{v}_{mn}) = G_v[\boldsymbol{0},\mathbf{I}], \text{ and} \tag{3}$$

$$Pr(\boldsymbol{x}_{mn}|\boldsymbol{u}_m,\boldsymbol{v}_{mn}) = G_x[\boldsymbol{\mu}+\mathbf{E}\boldsymbol{u}_m+\mathbf{A}\boldsymbol{v}_{mn},\Sigma], \tag{4}$$

where $G_a[\beta,\Gamma]$ is a Gaussian distribution in a with mean β and covariance Γ. Both the latent variables $\boldsymbol{u}_m$ and $\boldsymbol{v}_{mn}$ are specified with priors. In the original PLDA model, the objective of the learning is to estimate the parameters $\theta = \{\mathbf{E},\mathbf{A},\boldsymbol{\mu},\Sigma\}$, based on training data $\mathbf{X} = \{\boldsymbol{x}_{mn} \in R^d \,|\, m=1,...,M, n=1,...,N_m\}$, where d is the dimension of $\boldsymbol{x}$ and N_m is the number of images for identity m. Both the model parameters θ and latent variables are unknown. They can be jointly estimated by using the EM algorithm.

Before applying the EM algorithm, we can rewrite the model in (1) using matrix form:

$$\begin{aligned}\boldsymbol{x}_{mn} &= \boldsymbol{\mu}+[\mathbf{E}\ \ \mathbf{A}]\begin{bmatrix}\boldsymbol{u}_m\\\boldsymbol{v}_{mn}\end{bmatrix}+\boldsymbol{\varepsilon}_{mn}\\&= \boldsymbol{\mu}+\ \ \mathbf{B}\ \ \ \ \ \boldsymbol{z}_{mn}\ \ +\boldsymbol{\varepsilon}_{mn}\end{aligned} \tag{5}$$

In the E-step, the model statistics of the first two moments of z_{mn}, which is Gaussian distributed, are computed as follows [26]:

$$E[\boldsymbol{z}_m] = (\mathbf{B}^T\Sigma'^{-1}\mathbf{B}+\mathbf{I})^{-1}\mathbf{B}^T\Sigma'^{-1}(\boldsymbol{z}_m-\boldsymbol{\mu}'), \tag{6}$$

$$E[\boldsymbol{z}_m\boldsymbol{z}_m{}^T] = (\mathbf{B}^T\Sigma'^{-1}\mathbf{B}^T+\mathbf{I})^{-1}+E[\boldsymbol{z}_m]E[\boldsymbol{z}_m]^T, \tag{7}$$

where $\boldsymbol{\mu}' = [\boldsymbol{\mu}_m,\boldsymbol{\mu}_m,...,\boldsymbol{\mu}_m]^T$, and $\boldsymbol{\mu}_m$ is the mean face vector of the N_m images belonging to the same identity m, and Σ' is the diagonal matrix of Σ, where

$$\Sigma' = \begin{pmatrix}\Sigma & \cdots & 0\\\vdots & \ddots & \vdots\\0 & \cdots & \Sigma\end{pmatrix}. \tag{8}$$

In the M-step, we update the model parameters using the following rules:

$$\mu = \frac{1}{\sum_m N_m} \sum_m \sum_{N_m} x_{mn}, \tag{9}$$

$$\mathbf{B} = \left(\sum_{m,n} (x_{mn} - \mu) E[z_m]^T \right) \left(\sum_{m,n} E[z_m z_m^{\ T}] \right)^{-1}, \text{ and} \tag{10}$$

$$\Sigma = \frac{\sum_{m,n} \mathbf{Diag}\left[(x_{mn} - \mu)(x_{mn} - \mu)^T - \mathbf{B}E[z_m](x_{mn} - \mu)^T \right]}{\sum_m N_m}. \tag{11}$$

where **Diag**[ ] represents the operation of retaining only the diagonal elements from a matrix, and the updated **E** and **A** are computed from the new **B** using the equivalence between (6) and (7).

When looking at the model in (1) again, both **E** and **A** can be viewed as the corresponding identity and aging subspaces. Thus, solving the probabilistic problem in (4) can obtain the two subspaces at the same time [21]. In order to recognize face images accurately, the training samples are required to have both correct identity labels and aging labels at the same time; this imposes great difficulty in labelling the collected face samples. What's more, the desired identity and aging subspaces should be as independent of each other as possible, so as to separate the identity and aging factors as much as possible. However, using the identity and aging labels from the same dataset will inevitably lead to some correlation. These two practical problems push us to raise a bold question – can we obtain an ideal aging subspace, which is independent of the aging dataset being used? The answer is yes! We will show the details of establishing the aging subspace, and then compare it with the aging subspace, jointly learnt from the PLDA model in Section 2.2.

After obtaining the desired aging subspace, we then aim at optimizing the model in (1) by finding the parameters $\theta' = \{\mathbf{E}, \mu, \Sigma\}$ and the latent variables u_m and v_{mn}. As the aging subspace **A** is kept unchanged during the learning of the identity model, the EM algorithm can converge faster, with five to ten iterations only in our algorithm. We call this revised model as an identity-inference model. The algorithm for this part of model learning is summarized in Algorithm 1.

Algorithm 1 Identity-inference model learning

Input: The independent aging dataset $\mathbf{Z}$ with age group label $\mathbf{L}$, $\{z_l | l = 1, 2, ..., L\}$ based on appearance age and training dataset $\mathbf{X}$ with identity label $\mathbf{M}$, $\{x_m | m = 1, 2, ..., M\}$.

Output: Independent aging subspace $\mathbf{A}$ and dataset-specific identity subspace $\mathbf{E}$.

% Independent aging subspace learning %

1) Construct the adjacency graph within each age group using (12), where $w_{ij} = \exp(-\|z_i - z_j\|^2 / t)$ if i is among the k nearest neighbors of j in the same age group, or if j is among the K nearest neighbors of i in the same age group, otherwise $w_{ij} = 0$.

2) Compute the eigenvectors and eigenvalues for the generalized eigenvector problem in (13).

3) Compute the orthogonal basis vectors iteratively using (14) - (16), to form the orthogonal aging subspace $\mathbf{A}$ in (1), where $\mathbf{A} = [a_1, a_2, ..., a_k]$.

% Identity subspace learning %

1) Initialize the model parameter $\theta = \{\mathbf{E}, \mathbf{A}, \boldsymbol{\mu}, \Sigma\}$, where $\mathbf{E} = \text{rand}(*)$, $\mathbf{A} = [a_0, a_1, ..., a_{p-1}]$, $\boldsymbol{\mu} = \dfrac{1}{\sum_m N_m} \sum_{N_m} x_{mn}$ and $\sigma^2 = 0.1$.

Iterate:

2) E-step: update the latent variables z_m using (6) and (7).

3) M-step: update the model parameters $\theta = \{\mathbf{E}, \Sigma\}$ using (10) and (11).

Until convergence

2.2 Independent Aging Subspace Learning

As mentioned in the previous biometric studies [32], faces can be considered as points in a high-dimensional space, where aging is reflected by the distance of the face from the average of all face samples. It has been proved in [11] [33] that human aging effects can be projected onto a discriminant subspace using manifold learning, where it has a significant trend for sequential patterns. These works also used this subspace for image-based human age estimation on their own aging datasets, with ages ranged from 0 to 93 years. The findings enlighten us on the aging subspace learning with manifold, which can be later used in the identity-inference model. However, how to select a suitable aging face dataset and aging features for learning remains a difficult task.

2.2.1 Aging dataset for aging subspace learning

As mentioned before, the three most well-known aging datasets, FGNET, MORPH, and CACD, suffer from either lack of sufficient image data or lack of samples with time periods long enough for learning the sequential aging progression. Besides, all these datasets only have real-age labels, which are difficult for computers to learn, due to large individual variations as discussed before. All these facts make these datasets not ideal for aging subspace learning.

| Age 2 | Age 10 | Age 20 | Age 30 | Age 45 | Age 65 |

Fig. 2. Some face examples from the aging dataset Chalearn [35] with appearance age labels.

Recently, a new age-related dataset named Chalearn [35], whose face images are labelled with appearance ages, has been released. It is known to be the first dataset labeled with the appearance age instead of the real age. It contains 8,000 images, where the age of the face in each image was labeled by multiple individuals, and the average is taken as the appearance age. All the images are in the wild environments with real-life variations, and some of them are shown in Fig. 2. At the time we conducted our experiments, only the training set with 4,113 images of ages ranged from 1 to 86 years old had been released. Therefore, we used all these images, labeled with appearance ages, for the aging subspace learning.

2.2.2 Aging subspace learning

Manifold learning is one of the most widely used methods for data modeling in machine learning and can be applied to various applications like image classification [36], alignment [37] [38] and face analysis [39]. Discriminative locality [37] [40] is a critical property when applying manifold learning to obtain distribution nonlinearity of measurement and preserve discriminative information of data. It has been shown in [11] [33] that Locality Preserving Projection (LPP) [43] and its orthogonal variant OLPP [34] are able to project faces onto a more discriminative subspace, and characterize the age

manifold better than Principal Component Analysis (PCA) [44] and Locally Linear Embedding (LLE) [45]. Thus, OLPP is employed in our algorithm, which aims to preserve local structure based on the assumption that a nearest-neighbor search in the low-dimensional space will yield similar results to that in the high-dimensional space.

In the LPP theory, the objective function is defined as:

$$\sum_{ij}(z_i - z_j)^2 w_{ij}. \tag{12}$$

The weight w_{ij} is defined as $w_{ij} = \exp(-\|z_i - z_j\|^2 / t)$ when the two face features z_i and z_j are the k nearest neighbors of each other, otherwise $w_{ij} = 0$. The weight matrix $\mathbf{W}=[w_{ij}]$ is symmetric, a diagonal matrix $\mathbf{D}=[d_{ij}]$, whose entries are the column sums of $\mathbf{W}$, i.e. $d_{ii} = \sum_j w_{ij}$, and the corresponding Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$, can be computed. Then, the optimal projections can be obtained by solving the following eigenproblem:

$$\mathbf{ZLZ}^T a = \lambda \mathbf{ZDZ}^T a, \tag{13}$$

where the solutions are the column vectors $\{a_0,...,a_n\}$, which are the eigenvectors of $(\mathbf{ZDZ}^T)^{-1}\mathbf{ZLZ}^T$, with their eigenvalues in ascending order, i.e. $\lambda_0 < ... < \lambda_n$.

Table 1. The partitioning of ages into groups for the Chalearn dataset.

Group	Age range	#Image	Group	Age range	#Image
1	1-3	130	7	26-35	1222
2	4-6	147	8	35-45	602
3	7-10	96	9	46-55	367
4	11-15	89	10	56-60	142
5	16-20	396	11	61-65	71
6	21-25	772	12	66-89	79

In our algorithm, we apply OLPP to derive the orthogonal basis functions iteratively, which is denoted as $\{a_1, a_2,...,a_k\}$. We define:

$$\mathbf{A}^{(k-1)} = [\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k] \quad \text{and} \tag{14}$$

$$\mathbf{B}^{(k-1)} = [\mathbf{A}^{(k-1)}]^T (\mathbf{ZDZ}^T)^{-1} \mathbf{A}^{(k-1)}. \tag{15}$$

In this way, the orthogonal basis vectors $\{\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_k\}$ can be computed as: a. Compute $\boldsymbol{a}_1$ as the eigenvector of $(\mathbf{ZDZ}^T)^{-1}\mathbf{ZLZ}^T$ associated with the smallest eigenvalues; b. Compute $\boldsymbol{a}_k$ iteratively, as the eigenvector of

$$\mathbf{M}^{(k)} = \left\{ \mathbf{I} - (\mathbf{ZDZ}^T)^{-1} \mathbf{A}^{(k-1)} \cdot [\mathbf{B}^{(k-1)}]^{-1}[\mathbf{A}^{(k-1)}]^T \right\} \cdot (\mathbf{ZDZ}^T)^{-1} \mathbf{ZDZ}^T \tag{16}$$

Associated with the smallest eigenvalue of $\mathbf{M}^{(k)}$.

Since the appearance-age labels are given, we can further improve the learned manifold with supervised learning by utilizing the age-group label information. The weight w_{ij} is non-zero only for two face samples being the k nearest neighbors of each other and within the same age group, otherwise it is zero. The ages are partitioned into 12 groups, with the groups for the younger and older ages having smaller age intervals, as shown in Table 1.
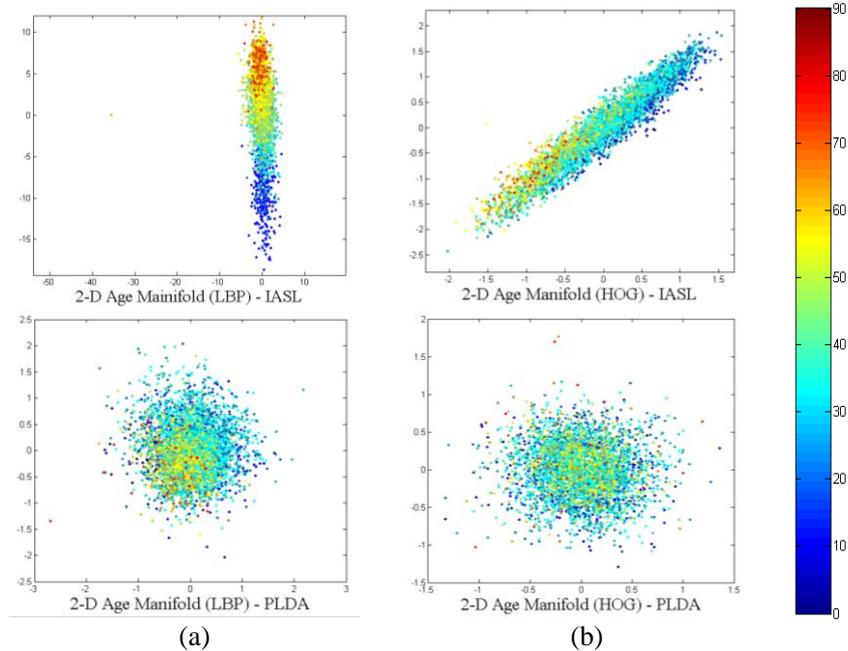

Fig. 3. 2-D age manifold visualization. (a) The two manifolds based on the wLBP features by using our Independent Aging Subspace Learning (IASL) and PLDA, respectively; and (b) the two manifolds based on the HOG features by using IASL and PLDA, respectively.

For each face image in the Chalearn dataset, we first locate the two eyes and align

the face, based on the eye positions as in [52]. We crop the faces to include the face regions only, and normalize them to the size 126×126 pixels. In order to alleviate the illumination impact, we normalize all the faces to have zero mean and unit variance. Previous work on age estimation [11] [33] used the whole face for feature extraction, which is not suitable for images under real-life variations. What's more, only using one feature is not sufficient for recognition tasks in the wild. In order to find the best features, supervised OLPP is applied to several state-of-the-art features, and the best two of these features, namely the multi-scale weighted Local Binary Patterns (wLBP) [46] and the Histogram of Oriented Gradients (HOG) [47] (specific configurations are given in the experimental session), are selected by using the same empirical mechanism as in [11]. As explained in [48], the LBP feature is able to capture the local texture information about a face, while the HOG feature represents the edge structure of a face well. In this sense, they are complementary to each other, and they are simple and fast to implement. We have studied the 2-D age manifolds on the different features to determine whether they can provide a distinct age progression. We have also applied PLDA on the Chalearn dataset to derive the aging subspace (where its basis are the columns of the matrix **D** in (1)).
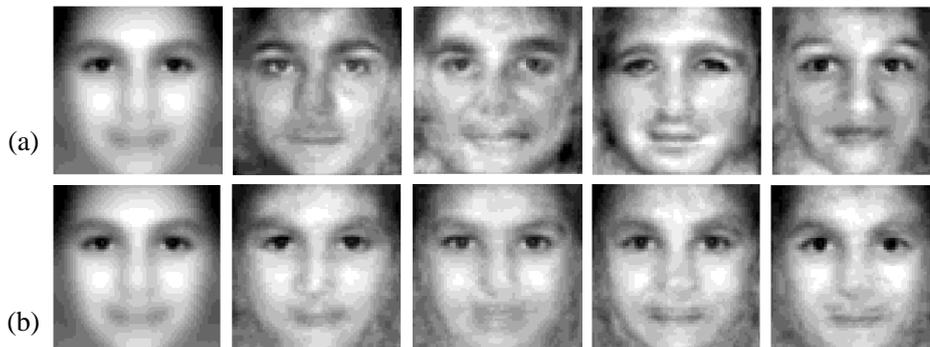

Fig. 4. Visualization of the identity inference model. (a) The mean face of all the face images and the faces in four directions in the identity subspace, where all the images look like different persons; and (b) the mean face and the faces in four directions in the aging subspace, where all images look like the same person but at different ages.

Fig. 3 illustrates the corresponding 2-D age manifolds, which shows that our aging subspaces learnt from the two features have much more distinctive patterns of aging progression than those obtained from the PLDA model. Fig. 4 gives further

visualization of how the identity subspace captures the faces with different appearances (E varies while A stays constant) and how the aging subspace illustrates the faces with different ages (A varies while E stays constant). More experiment results are given in Section 4.

3. Face Recognition Based on Aging-Guided Identity-Inference Model

In this section, the overall framework of the age-invariant face recognition, based on the proposed identity-inference model, is presented. Fig. 5 shows the overall framework, which includes pre-processing steps, such as feature extraction and dimension reduction, face recognition after obtaining the identity subspace, and face matching based on different feature-fusion schemes. For the convenience of readers, the important notations used in this section have been summarized in Table 2.

Table 2. The important notations used in Section 3.

Notation	Description	Group	Description
$\mathbf{A}$	aging subspace	$\mathbf{F}_{HOG}$	HOG feature
$\mathbf{B}$	identity subspace	α	direction matrix for $\mathbf{F}_{wLBP}$
$\boldsymbol{x}_p$	input probe face image	β	direction matrix for $\mathbf{F}_{HOG}$
$\boldsymbol{x}_M$	gallery face image	$\boldsymbol{g}_{wLBP}$	projected weighted LBP feature
$\mathbf{F}_{wLBP}$	weighted LBP feature	$\boldsymbol{g}_{HOG}$	projected HOG feature

3.1 Feature Extraction

As mentioned in previous section, local features, such as LBP and HOG, have been proven to have more discriminative power and are widely used in face recognition. Furthermore, they are easy and fast to implement. In our independent age subspace learning, we have also found that these two local features could achieve the best performance in terms of aging progression representation, as illustrated in Fig. 3. Thus, we use both wLBP and HOG as the feature descriptors in our experiments throughout the paper.
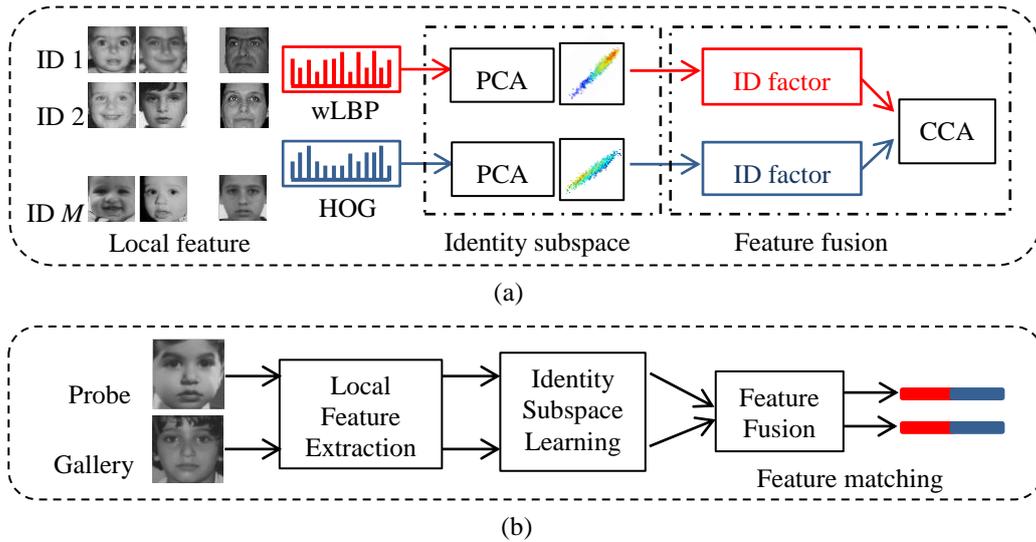
Fig. 5. The overall framework of the proposed age-invariant face recognition algorithm based on identity inference with independent aging subspace learning: (a) the training phase and (b) the recognition phase.

For all the face images, we perform the same preprocessing steps as described in Section 2.2.2. The wLBP features are extracted with 7×7 windows, at three different radii {1, 3, 5} (due to the limited size of images), which has been found to achieve the best performance. As each of the windows has a different degree of importance, different weights are assigned to them, as illustrated in Fig. 6.
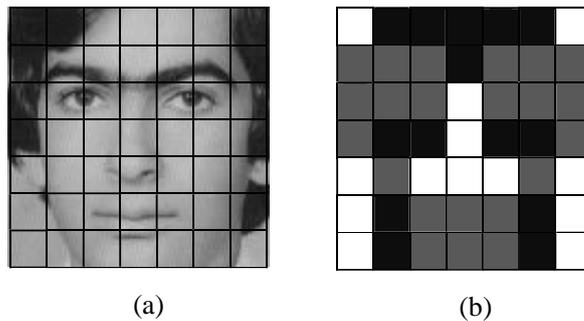

Fig. 6. (a) A cropped face partitioned into 7×7 windows for extracting the weighted MLBP features, and (b) the weights used for the MLBP features in each partition, where black, gray, and white represent the weights of 3, 2, and 1, respectively.

Unlike the previous LBP features designed for face recognition [46] and face retrieval tasks [48], the weight mask used in our framework places greater importance on those regions that are more easily influenced by aging effects, such as the forehead, cheeks, and mouth corners [24]. For the HOG feature, by experiment, the best

performance can be achieved if the patch size is 21×21 pixels, with an overlapping factor of 0.5 and 4 orientations. In order to facilitate dimensionality reduction for a small-size dataset, like FGNET, we use all the face images in the Chalearn dataset to determine the PCA subspace, with 95% variances retained.

3.2 Face Recognition Based on AG-IIM

By applying the supervised OLPP, we can independently learn a discriminative aging subspace that can better represent the aging progression. After obtaining the aging subspace $\mathbf{A}$ of the model shown in (5), we plug it into the model for the computation of the identity subspace $\mathbf{E}$. It should be noticed that the composite matrix $\mathbf{B} = [\mathbf{E} \ \mathbf{A}]$ needs to be updated as a whole in order to optimize the model statistics $E[z_m]$ and $E[z_m z_m^T]$. Our experiments show that fixing the aging subspace $\mathbf{A}$ can produce a better identity subspace $\mathbf{E}$ for face recognition. This is due to the fact that our proposed aging subspace is learnt by using face images with appearance ages. Thus, in the matrix $\mathbf{B}$, only the identity subspace is updated, i.e. $\mathbf{B}' = [\mathbf{E}' \ \mathbf{A}]$.

In the recognition stage, we calculate the predictive distributions between the input probe image $\boldsymbol{x}_p$ and each gallery image, i.e. $p_r(\boldsymbol{x}_p|\boldsymbol{x}_1)$, $p_r(\boldsymbol{x}_p|\boldsymbol{x}_2)$, ..., and $p_r(\boldsymbol{x}_p|\boldsymbol{x}_M)$, and then evaluate the likelihood for each of these distributions. As proved in [28], the likelihood between the probe and a gallery image takes in a Gaussian form, and can be simplified by projecting the data onto a subspace similar to the original LDA method. Suppose that the probe image $\boldsymbol{x}_p$ is put into the identity-inference model in (5), the output feature vector can then be computed as follows:

$$
\begin{aligned}
\boldsymbol{f}_p &= \mathbf{E}^T (\mathbf{B}\mathbf{B}^T + \boldsymbol{\Sigma}')^{-1}(\boldsymbol{x}_p - \boldsymbol{\mu}) \\
&= \mathbf{E}^T (\mathbf{E}\mathbf{E}^T + \mathbf{A}\mathbf{A}^T + \boldsymbol{\Sigma}')^{-1}(\boldsymbol{x}_p - \boldsymbol{\mu}).
\end{aligned}
\tag{17}
$$

In this way, it is equivalent to projecting the input probe image into a discriminative subspace, where a distance metric can be used for face recognition.

3.3 Feature Fusion Scheme for Face Matching

In order to further improve the face recognition performance, both the wLBP and HOG features are used and fused in our framework. One way to perform feature fusion is to compute the *z*-score, where both feature vectors are normalized and then concatenated to form a long feature vector. This is the simplest way, but does not take the correlation between the two features into consideration. What's more, concatenating two different types of features directly may cancel out their discriminative power, which leads to an even lower recognition rate. As the two features are extracted from the same identity, they should be correlated. Therefore, as in [53], CCA is used to project the two features into a coherent subspace, where the correlation between them is maximized. The projected features are then combined to form a single coherent feature vector for age-invariant face recognition.

With the pairs of output features computed with the identity-inference model (17), denoted as $\mathbf{F}_{wLBP}$ and $\mathbf{F}_{HOG}$, we apply CCA to learn the pairs of directions α and β that maximize the correlation between the projected features, i.e. $g_{wLBP} = \alpha^T f_{wLBP}$ and $g_{HOG} = \beta^T f_{HOG}$, with the correlation between g_{wLBP} and g_{HOG} maximized. The direction matrices α and β can be derived by maximizing the following criterion function:

$$\rho = \frac{E[g_{wLBP} g_{HOG}]}{\sqrt{E[g_{wLBP}^2] E[g_{HOG}^2]}} = \frac{\alpha^T \mathbf{C}_{12} \beta}{\sqrt{\alpha^T \mathbf{C}_{11} \alpha \cdot \beta^T \mathbf{C}_{22} \beta}}, \tag{18}$$

where $\mathbf{C}_{11}$ and $\mathbf{C}_{22}$ denote the covariance matrices of g_{wLBP} and g_{HOG}, respectively, and $\mathbf{C}_{12}$ is the covariance matrix of g_{wLBP} and g_{HOG}.

In the training stage, after $\mathbf{F}_{wLBP}$ and $\mathbf{F}_{HOG}$, whose columns are the feature vectors f_{wLBP} and f_{HOG}, are computed by using (17), we normalize them to have zero mean and unit variance. Then, the projection matrices α and β are computed by using (18). It can be shown that the optimal direction matrices α and β are the eigenvectors of $\mathbf{R}_1 = \mathbf{C}_{11}^{-1} \mathbf{C}_{12} \mathbf{C}_{22}^{-1} \mathbf{C}_{21}$ and $\mathbf{R}_2 = \mathbf{C}_{22}^{-1} \mathbf{C}_{21} \mathbf{C}_{11}^{-1} \mathbf{C}_{12}$, respectively. When the

pair of features, $\boldsymbol{f}_{p\text{-}wLBP}$ and $\boldsymbol{f}_{p\text{-}HOG}$, of an input probe image are obtained in the testing stage, we further project them into the corresponding CCA subspaces to form coherent features, as follows:

$$\boldsymbol{g}_p = [\boldsymbol{g}_{p\text{-}wLBP} \; \boldsymbol{g}_{p\text{-}HOG}], \tag{19}$$

where $\boldsymbol{g}_{p\text{-}wLBP} = \boldsymbol{\alpha}^T \boldsymbol{f}_{p\text{-}wLBP}$ and $\boldsymbol{g}_{p\text{-}HOG} = \boldsymbol{\beta}^T \boldsymbol{f}_{p\text{-}HOG}$. Then, the Euclidean distance is computed, and the nearest-neighbor rule is used for face recognition.

4. Experiments and Analysis

To evaluate the performance of our proposed aging-guided identity-inference model (AG-IIM) for age-invariant face recognition, we compare it with several state-of-the-art methods on different datasets, namely the FGNET dataset [6], the MORPH dataset [30], and the CACD dataset [31]. FGNET is known to be the first popular face aging dataset and has been widely used for evaluating age-related facial image analysis tasks. It contains 1,002 images of 82 individuals, and the images were collected at ages ranging from 0 to 69. The MORPH dataset was proposed later, and contains a larger number of subjects. It has two sections, namely MORPH album one and MORPH album two. As album one is small (only 1,690 face images in total), most recent works use album two for experiments, as it has 55,134 facial images of 13,617 persons. The CACD dataset is the latest aging dataset, which contains 163,446 images of 2,000 celebrity individuals retrieved from the Internet. Some statistics and sample facial images are given in Table 3 and Fig. 7, respectively. It can be seen that FGNET is the most challenging dataset, as it has the smallest number of images but the largest age gap, while all the photos are also taken under large variations.

Table 3. Statistics of the face aging datasets.

Dataset	#Image	#Identity	Age range	Age gap	In wild
FGNET	1,002	82	0-69	0-45	Yes
MORPH	55,134	13,617	16-77	0-5	No
CACD	163,446	2,000	16-62	0-10	Yes

Fig. 7. Sample face images from the comparison datasets: (a) FGNET dataset, (b) MORPH dataset and (c) CACD dataset.

4.1 Face Recognition on the FGNET Dataset

To fully evaluate our proposed model, we first present the parameters for the wLBP and HOG features used in feature extraction, the aging subspace learning, and the identity inference, in Table 4. All these parameters are empirically chosen with reference to previous related works and our own experiment results. It is noticeable that both wLBP and HOG features perform the best with similar dimension representation and nearest neighbors to learn aging subspace with OLPP. However, it seems that wLBP feature needs more eigenvectors to capture aging information while HOG needs more to capture identity information.

Table 4. Parameter settings for FGNET dataset.

Parameters		wLBP	HOG
Feature Configuration	PCA variance	95%	95%
	Feature dimension	1473	1371
Aging Learning Model (Supervised OLPP)	# Nearest neighbors	6	5
	Heat kernel coef. t	1	1
Identity Inference Model in (1)	#Aging eigenvectors	600	150
	#Identity eigenvectors	80	350

One of the biggest advantages of our proposed method is that, during experiments, age labels of training samples are no longer needed, because we have independently learnt the aging subspace for the identity-inference model. Furthermore, as the FGNET dataset only has 1,002 images in total, while the feature dimensions are much higher, we have applied a simple but effective approach to solving the overfitting problem. Unlike the previous strategies [20] [21], which applied random subspaces and feature slicing, we use the ChaLearn images, together with the FGNET images, to learn the PCA subspace, with 95% of the variance retained. In this way, the more discriminative power of the training features can be preserved, while projecting to the same PCA subspace with aging images from the ChaLearn dataset also improves the aging pattern learning.

Our algorithm is fine-tuned using the FGNET dataset. We will later show that the model, trained by using FGNET, can also be applied to recognize faces from other datasets. The performances are similar irrespective of whether or not the training and testing images are from the same dataset. We conducted a thorough evaluation and comparison with some recent, state-of-the-art age-invariant face recognition methods. These include: (a). one of the earliest frameworks on age-invariant face recognition [19], which establishes a 3D aging modeling scheme for age correction for recognition; (b). a discriminative model proposed in [20]; (c). a hidden factor analysis framework [21], which separates aging and identity at the same time; (d). a feature-aging model, which uses local Gabor feature and linear mapping to predict the aging of facial features [49]; and (e). a two-step framework [5], based on a maximum entropy feature descriptor, and identity factor analysis matching, which has achieved the best recognition performance on FGNET previously. Following the same experiment set-up, we evaluate all these methods, in leave-one-person-out fashion, based on the rank-1 recognition rates, as shown in Table 4. It should be noted that, except [49], all the results of the compared methods, reported in this paper, are based on the best results as reported in their respective papers.

From Table 5, we can see that the proposed AG-IIM, based on independent aging subspace learning, achieves better performances than other methods. Furthermore, the HOG feature is more effective than the LBP feature for representing the aging factors.

By fusing the two features using CCA, the recognition rate can be further improved significantly. To the best of our knowledge, this is the highest recognition accuracy that has ever been achieved on the FGNET dataset, which is known to be the most challenging dataset for age-invariant face recognition.

Table 5. Rank-1 recognition rates on the FGNET dataset.

Algorithms	Recognition Rates
3D aging model (2010) [19]	37.4%
Discriminative aging model (2011) [20]	47.5%
Hidden factor analysis model (2013) [21]	69.0%
Feature-aging model (2015) [49]	71.3%
Maximum entropy model (2015) [5]	76.2%
AG-IIM with the MLBP feature only	**80.8%**
AG-IIM with the HOG feature only	**84.14%**
AG-IIM with feature fusion by CCA	**88.23%**

(a)

| Age: 1 | Age: 29 | Age: 34 | Age: 3 | Age: 22 | Age: 12 |

(b)

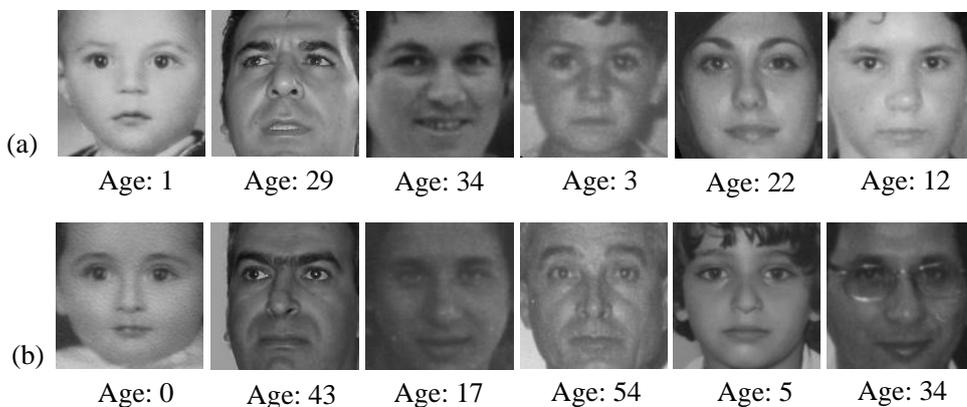| Age: 0 | Age: 43 | Age: 17 | Age: 54 | Age: 5 | Age: 34 |

Fig. 8. Some examples of the recognition failure cases on the FGNET dataset: (a) input query images, and (b) the corresponding retrieved images, based on Euclidean distance.

In the experiments, we have also studied some of the failure cases, in which our proposed AG-IIM could not find the correct subject in the gallery. Among most of these cases, the failure is mainly due to the similar appearances of the query and the gallery faces, at similar ages (appearance ages), as shown in the first three columns in Fig. 8. We have also observed some interesting results, for which the input query images were matched to training images with a large age difference, as shown in the last three

columns in Fig. 8. After a close-up analysis, some underlying similarities between these incorrect retrieved pairs can be observed, such as noses, mouths, facial structures, etc. This further shows that the proposed framework is seeking the substantial identity information, instead of being tricked by the superficial likeness.

4.2 Face Recognition on the MORPH Dataset

In this section, we extend the experiment on the MORPH dataset to examine the efficiency of the proposed method. Unlike FGNET, the MORPH dataset does not have a large age gap for each of its subjects. However, all its face images are still under large pose, lighting, and expression variations. For this dataset, we follow the same split rule, as the previous method [20], where 10,000 individuals are randomly selected. Then, the youngest face images of the selected subjects are used to form the gallery set, while the corresponding oldest images are used to form the probe set. In this way, both the gallery and probe sets have 10,000 images from the different individuals. All the images of the remaining 3,617 subjects are used for training the identity subspace based on the identity-inference model, where the aging subspace is still learnt independently. Same as FGNET, only the identity labels of the training images are used during the experiment, which makes the training process much simpler.

For the proposed AG-IIM method, the same pre-processing steps, including eye detection, face alignment, and feature extraction, are applied to all the images in the MORPH dataset, as described in the previous section. Furthermore, in addition to using the MORPH images for training, we also use the identity subspace learnt from the FGNET dataset, i.e. the identity subspace is learnt without using the MORPH database, so that the generalization ability of our proposed algorithm can be tested. To fully evaluate the recognition performance of the proposed algorithm, we compare the rank-1 recognition rate with some of the state-of-the-art methods: (a). the 3D aging model [19]; (b). the discriminative aging model [20]; (c). the hidden factor analysis model (2013) [21]; (d). the maximum entropy model (2015) [5]; (e). the cross-age reference coding (CARC) model, and (f). local pattern selection with the hidden factor analysis (LPS+HFA) model [50], which is an extension of [5] and has achieved the best

recognition accuracy on MORPH, to date. The comparison results of the different methods are compared and shown in Table 6.

Table 6. Rank-1 recognition accuracies on the MORPH dataset.

Algorithms	Recognition Accuracies
3D aging model (2010) [19]	79.8%
Discriminative aging model (2011) [20]	83.9%
Hidden factor analysis model (2013) [21]	91.14%
Maximum entropy model (2015) [5]	92.26%
CARC model (2015) [31]	92.8%
LPS+HFA model (2016) [50]	94.87%
AG-IIM trained on FGNET	**93.12%**
AG-IIM trained on MORPH	**95.62%**

As some existing methods have already achieved impressive performance results on the MORPH dataset, the improvement of our method in terms of recognition accuracy is marginal, but our method is still comparable to the state-of-the-art methods. More importantly, we found that, even if we replace the training data with the 1,002 images from FGNET and with their identity labels, a similar recognition performance can still be obtained. This finding is exciting in the sense that, from Fig. 7, most of the images from the FGNET and the MORPH datasets have different races. However, because the subjects in the FGNET dataset have a wide range of age difference, our algorithm can capture more substantial identity information, which compensates for the difficulties of recognition across races.

4.3 Face Verification on the CACD Dataset

In order to fully examine the generalization power of our proposed framework for age-invariant face recognition, we further conducted face verification on the CACD verification subset (CACD-VS). It comes as a part of the CACD dataset, and contains 2,000 positive pairs (images of the same person across ages) and 2,000 negative pairs. They are carefully selected and annotated to make sure that each of the images has a

correct identity tag. To perform verification, we use the marginalized likelihood [28] as the metric learning for recognition. For each pair of face images, we compute the likelihood $p_r(x_p, x_g)$ that they belong to the same identity, and the likelihood $p_r(x_p)p_r(x_g)$ that they are from different identities. Then, we compute the likelihood ratio, and compare it to a threshold for face verification, as in [28]:

$$R(x_p, x_g) = \frac{likelihood(same)}{likelihood(diff)} = \frac{p_r(x_p, x_g)}{p_r(x_p)p_r(x_g)} . \tag{20}$$

Following the same experiment set-up in [31], we partition the whole subset into ten folds, with 400 image pairs (200 positive and 200 negative) in each fold. In order to test the generalization power of the proposed framework, we still use all the FGNET images for training to obtain the identity subspace. What's more, we notice that, as all the subjects in CACD are celebrities retrieved from online, most of the women's faces are wearing make-up, which makes the recognition across ages become more difficult. To further assist the model in adapting different variations, we also added 798 face images of 100 identities (which were chosen from one of the folders and then excluded from testing) from the CACD dataset. In this way, the training set size is increased to 1800 images, in total. The threshold for (20) is learnt by using 3,200 pairs of images from eight of the folds, while the remaining single fold serves as the testing set. The experiments are repeated nine times, and the average verification results are shown in Table 7. The performance of our method is compared to several state-of-the-art methods, as well as the human voting reported in [31], in terms of the receiver-operating characteristic (ROC) curves, as shown in Fig. 9.

From the results, the proposed method outperforms other methods and is better than the average performance of humans, although the number of training samples used in our method is fewer than the methods reported in [31]. However, the verification performance of our method is relatively low on the CACD-VS dataset, as the images have more variations besides age. It should also be noted that the voting-based human performance is still better than machine performance, especially when the faces not only have large age variations, but also variations in poses, expressions, illuminations, etc. A

possible future work for our proposed method can investigate age-invariant face recognition under real-life challenges.

Table 7. Verification accuracies on the CACD-VS dataset.

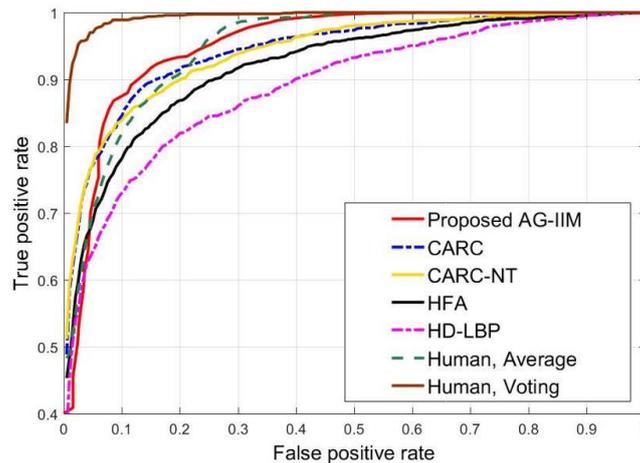| Algorithms	Verification Accuracies
High Dimensional LBP (2013) [51]	81.6%
Hidden factor analysis model (2013) [21]	84.4%
CARC-NT model (2015) [31]	85.6%
CARC model (2015) [31]	87.6%
AG-IIM trained on FGNET	**89.8%**
Human, Average	85.7%
Human, Voting	94.2%

Fig. 9. ROC curves for verification task on CACD-VS dataset.

4.4 Comparison of the use of real age and appearance age

In order to verify the efficiency of using appearance age and independent learning of age subspace based on PLDA, we conducted controlled experiments on the three abovementioned aging datasets.

Besides our proposed AG-IIM, we also use the real-age labels from the FGNET, MORPH, and CACD datasets. We compare and evaluate using real age and appearance age for learning the identity subspace in our proposed model. The algorithm based on real age is denoted as real-aging-guided identity-inference model (rAG-IIM). We also further test the efficiency of independent aging subspace learning using real age labels

and the PLDA model to jointly derive the aging and identity subspaces. This method is denoted as real-aging identity-inference model (rA-IIM). We apply face recognition/verification and summarize the performances of the three methods in Table 8.

Table 8. Verification accuracies of different models.

Datasets / Models	FGNET	MORPH	CACD
AG-IIM	88.2%	95.6%	89.9%
rA-IIM	69.9%	90.8%	84.6%
rAG-IIM	70.1%	91.2%	88.7%

It is obvious that using appearance-age labels and independently learning the aging and identity subspaces can achieve the best verification accuracy for all the aging datasets. This also shows the potential of using an independent aging dataset with appearance-age labels for the age-invariant face recognition task, where we no longer need the real age labels from the experiment datasets. This approach can greatly alleviate the issue, where there are large variances for real-age faces, like those in FGNET, and there are only a few age differences among the faces collected within a short period, like those in CACD.

5. Conclusions and Future Work

In this paper, we have proposed an aging-guided identity-inference model (AG-IIM), based on independent aging subspace learning for age-invariant face recognition. Following the idea that a face can be separated into an identity representation, which is invariant to the aging influence, and an aging representation, which changes with age progression, we have applied the Probabilistic LDA (PLDA) model to obtain two distinctive subspaces. Unlike all the previous methods, which utilize real-age labels, we have proposed using appearance-age labels to learn the aging subspace independently. It has been shown that, using appearance-age information can help computers better understand the relationship between ages and the corresponding

facial appearances. The performance of our proposed method on age-invariant face recognition can be improved. What's more, learning the aging subspace independently enables experiments to rely only on the identity labels given by the datasets, which makes constructing a face-aging dataset much easier. Another contribution of this paper is using the projection of two efficient feature representations onto a more correlated subspace by using Canonical Correlation Analysis (CCA), where their correlation is maximized. Experimental results on different datasets have shown superior performances for our method in terms of recognition accuracy, especially on the most challenging aging dataset FGNET, where face images have the largest age range.

As mentioned in experiment section, there is still a long way to go before computer-based age-invariant face recognition will surpass the human performance. Our future work will focus on the further improvements of the identity-inference model. More statistical analysis and constraints may be applied to make it more robust against variations other than age.

References

[1] Y. Sun, D. Liang, X. Wang and X. Tang, "Deepid3: Face recognition with very deep neural networks," *arXiv preprint,* vol. 1502, no. 00873, 2015.

[2] F. Schroff, D. Kalenichenko and J. Philbin, "Facenet: A unified embedding for face recognition and clustering.," in *in IEEE Int, Conf. Computer Vision and Pattern Recognition*, 2015.

[3] C. Ding and D. Tao, "Robust face recognition via multimodal deep face representation," *IEEE Trans.Multimedia,* vol. 17, no. 11, pp. 2049-2058, 2015.

[4] G.B.Huang, M. Ramesh, T. Berg and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," *Technical Report,* vol. 1, no. 2, p. 3, 2007.

[5] D. Gong, Z. Li, D. Tao, J. Liu and X. Li, "A Maximum Entropy Feature Descriptor for Age Invariant Face Recognition," in *IEEE. Int. Conf. Computer Vision and Pattern Recognition*, 2015.

[6] T. Cootes and A. Lanitis, "FG-NET Aging Database," 2008.

[7] H. Ling, S. Soatto, N. Ramanathan and D. Jacobs, "A study of face recognition as people age," in *presented at IEEE Int. Conf. Computer Vision*, 2007.

[8] A. Lanitis, "A survey of the effects of aging on biometric identity verification," *Int. J. Biometrics,* vol. 2, no. 1, pp. 34-52, 2010.

[9] Y. Fu, G. Guo and T. S. Huang, "Age synthesis and estimation via faces: A survey," *IEEE Trans, Pattern Anal. Mach. Intell.,* vol. 21, no. 11, pp. 1955-1976, 2010.

[10] X. Geng, Z. H. Zhou and K. Smith-Miles, "Automatic age estimation based on facial aging patterns," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 29, no. 12, pp. 2234-2240, 2007.

[11] Y. Fu and T. S. Huang, "Human age estimation with regression on discriminative aging manifold," *IEEE Trans. Multimedia,* vol. 10, no. 4, pp. 578-584, 2008.

[12] G. Mu, G. Guo, Y. Fu and T. S. Huang, "Human age estimation using bio-inspired features," in *IEEE Conf. Comput. Vis. Pattern Recog.*, 2009.

[13] A. Montillo and H. Ling, "Age regression from faces using random forests," in *IEEE Int. Conf. Image Process.*, 2009.

[14] K. Zhu, D. Gong, Z. Li and X. Tang, "Orthogonal gaussian process for automatic age estimation," in *ACM Int. Conf. Multimedia*, 2014.

[15] A. Lanitis, C. J. Taylor and T. F. Cootes, "Toward automatic simulation of aging effects on face images," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 24, no. 4, pp. 442-455, 2002.

[16] J. Suo, S. C. Zhu, S. Shan and X. Chen, "A compositional and dynamic model for face aging," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 32, no. 3, pp. 385-401, 2010.

[17] J. Suo, X. Chen, S. Shan and W. Gao, "Learning long term face aging patterns from partially dense aging databases," in *IEEE Int. Conf. Computer Vision*, 2009.

[18] X. Shu, J. Tang, H. Lai, L. Liu and S. Yan, "Personalized Age Progression with Aging Dictionary," in *IEEE Int. Conf. Computer Vision*, 2015.

[19] U. Park, Y. Tong and A. K. Jain, "Age-invariant face recognition," *IEEE Trans. Pattern Anal.*

Mach. Intell., vol. 32, no. 5, pp. 947-954, 2010.

[20] Z. Li, U. Park and A. K. Jain, "A discriminative model for age invariant face recognition," *IEEE Trans. Inf. Forens. Security,* vol. 6, no. 3-2, pp. 1028-1037, 2011.

[21] D. Gong, Z. Li, D. Lin, J. Liu and X. Tang, "Hidden factor analysis for age invariant face recognition," in *IEEE Int. Conf. Computer Vision*, 2013.

[22] B. Chen, C. Chen and W. H. Hsu, "Cross-age reference coding for age-invariant face recognition and retrieval," in *European Conf. on Computer Vision*, 2014.

[23] H. Ling, S. Soatto, N. Ramanathan and D. W. Jacobs, "Face verification across age progression using discriminative methods," *IEEE Trans. Inf. Forens. Security,* vol. 5, no. 1, pp. 82-91, 2010.

[24] C. Otto, H. Han and A. K. Jain, "How does aging affect facial components?," in *European Conf. on Computer Vision*, 2012.

[25] L. Du and H. Ling, "Cross-age face verification by coordinating with cross-face age verification," in *IEEE Conf. Comput. Vis. Pattern Recog.*, 2015.

[26] S. J. Prince and J. H. Elder, "Probabilistic linear discriminant analysis for inferences about identity," in *IEEE Int. Conf. Computer Vision*, 2007.

[27] S. Ioffe, "Probabilistic linear discriminant analysis," in *IEEE Int. Conf. European Conf. on Computer Vision*, 2006.

[28] P. Li, Y. Fu, U. Mohammed, J. H. Elde and S. J. Prince, "Probabilistic models for inference about identity," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 34, no. 1, pp. 144-157, 2012.

[29] A. Dempster, N. Laird and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the royal statistical society,* pp. 1-38, 1977.

[30] K. Ricanek and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," in *IEEE Int. Conf. Automat. Face Gesture Recog.*, 2006.

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

[32] K. A. Deffenbacher, T. Vetter, J. Johanson and O'Toole, "Facial aging, attractiveness, and aistinctiveness," *Perception,* vol. 27, pp. 1233-1243, 1998.

[33] G. Guo, Y. Fu, C. Dyer and T. S. Huang, "Image-based human age estimation by manifold learning and locally adjusted robust regression," *IEEE Trans. Image Processing,* vol. 17, no. 7, pp. 1178-1188, 2008.

[34] D. Cai, X. He, J. W. Han and H. Zhang, "Orthogonal laplacianfaces for face recognition," *IEEE Trans. Image Processing,* vol. 15, no. 11, pp. 3608-3614, 2006.

[35] "ChaLearn Looking at People and Faces of the World 2016," online: http://gesture.chalearn.org/.

[36] J. Yu, D. Tao and M. Wang, "Adaptive hypergraph learning and its application in image classification," *IEEE Tans. Image Processing*, vol. 21, no. 7, pp. 3262-3272, 2012.

[37] T. Zhang, D. Tao, X. Li and J. Yang, "Patch alignment for dimensionality reduction", *IEEE Trans. Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1299-1313, 2009.

[38] J. Yu, R. Hong, M. Wang and J. You, "Image clustering based on sparse patch alignment framework," *Pattern Recognition*, vol. 47, no. 11, pp. 3512-3519, 2014

[39] C. Hong, J. Yu and D. Tao, "Multimodal deep autoencoder for human pose recovery." *IEEE Trans. Image Processing*, vol. 24, no. 12, pp. 5659-5670, 2015.

[40] C. Hong, J. Yu, J. You, C. Chen and D. Tao, "Multi-view ensemble manifold regularization for 3D object recognition", *Information Sciences*, vol. 1, no. 320, pp. 395-405, 2015.

[41] D. R. Hardoon, S. Szedmak and J. Shawe-Taylor, "Canonical correlation analysis: An overview with application to learning methods," *Neural computation,* vol. 16, no. 12, pp. 2639-2664, 2004.

[42] Y. J. Zhang, Ed., Advances in Face Image Analysis: Techniques and Technologies, IGI Global, 2010.

[43] X. Niyogi, "Locality preserving projections," *Neural information processing systems,* vol. 16, p. 153, 2004.

[44] P. Belhumeur, J. Hespanha and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 19, no. 7, pp. 711-720, 1997.

[45] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science,* vol. 290, no. 5500, pp. 2323-2326, 2000.

[46] T. Ahonen, A. Hadid and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 28, no. 12, pp. 2037-2041, 2006.

[47] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Conf. Comput. Vis. Pattern Recog.*, 2005.

[48] H. Zhou, K. Lam and X. He, "Shape-appearance-correlated active appearance model," *Pattern Recognition,* vol. 56, pp. 88-99, 2016.

[49] H. Zhou, K. Wong and K. Lam, "Feature-aging for age-invariant face recognition," in *IEEE Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2015.

[50] Z. Li, D. Gong, X. Li and D. Tao, "Aging Face Recognition: A Hierarchical Learning Model Based on Local Patterns Selection," *IEEE Trans. Image Processing,* vol. 25, no. 5, pp. 2146 - 2154, 2016.

[51] D. Chen, X. Cao, F. Wen and J. Sun, "Blessing of dimensionality: High-dimensional feature and its efficient compression for face verification," in *IEEE Conf. Compuat. Vis. Pattern Recog.*, 2013.

[52] K. M. Lam and H. Yan, "Locating and Extracting the Eye in Human Face Images", *Pattern Recognition*, vol. 29, no. 5, pp. 771-9, 1996.

[53] K. H. Pong and K. M. Lam, "Multi-resolution feature fusion for face recognition", *Pattern Recognitio*n, vol. 47, no. 2, pp. 556-567, 2014.