

# Hidden Factor Analysis for Age Invariant Face Recognition

Dihong Gong<sup>1</sup>  
dh.gong@siat.ac.cn

Zhifeng Li<sup>1</sup>  
zhifeng.li@siat.ac.cn

Dahua Lin<sup>2</sup>  
dhlin@ttic.edu

Jianzhuang Liu<sup>3,4</sup>  
liu.jianzhuang@huawei.com

Xiaoou Tang<sup>1,3</sup>  
xtang@ie.cuhk.edu.hk

<sup>1</sup>Shenzhen Key Lab of Computer Vision and Pattern Recognition  
Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

<sup>2</sup>Toyota Technological Institute at Chicago

<sup>3</sup>Dept. of Information Engineering, The Chinese University of Hong Kong

<sup>4</sup>Media Lab, Huawei Technologies Co. Ltd., China

## Abstract

Age invariant face recognition has received increasing attention due to its great potential in real world applications. In spite of the great progress in face recognition techniques, reliably recognizing faces across ages remains a difficult task. The facial appearance of a person changes substantially over time, resulting in significant intra-class variations. Hence, the key to tackle this problem is to separate the variation caused by aging from the person-specific features that are stable. Specifically, we propose a new method, called Hidden Factor Analysis (HFA). This method captures the intuition above through a probabilistic model with two latent factors: an identity factor that is age-invariant and an age factor affected by the aging process. Then, the observed appearance can be modeled as a combination of the components generated based on these factors. We also develop a learning algorithm that jointly estimates the latent factors and the model parameters using an EM procedure. Extensive experiments on two well-known public domain face aging datasets: MORPH (the largest public face aging database) and FGNET, clearly show that the proposed method achieves notable improvement over state-of-the-art algorithms.

## 1. Introduction

As an emerging research topic, age invariant face recognition has many practical applications. For example, in law enforcement, finding missing children or identifying criminals based on their mug shots on identity requires recognizing photos across ages [3, 29]. In spite of the great advancement in face recognition in the past decades, age invariant

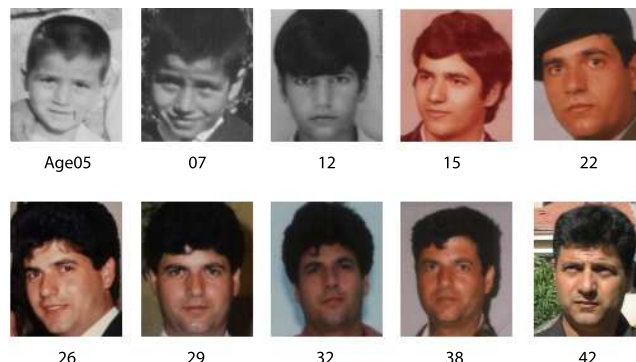


Figure 1. Example images showing the large intra-class variations due to facial aging for one of the subjects in the FG-NET database [2].

face recognition remains as a major challenge. The difficulty of this problem, to a great extent, arises from the fact that the face appearance of a person is subject to remarkable change caused by the aging process over time, as shown in Figure 1.

The research on age related face image analysis has only been studied in recent years. Most existing works focus on age estimation [8, 10–12, 17, 18, 24, 28, 37, 40, 41] and aging simulation [7, 19, 27, 30, 31, 35]. However, work that explicitly tackles age invariant face recognition is limited. Existing methods on age invariant face recognition roughly fall into two categories: generative approaches and the discriminative approaches. Generative methods try to synthesize face images that match the target age before recognition [7, 10, 19, 27]. They try to construct a 2-D or 3-D generative model to compensate for the aging process in face matching. These methods, however, typically suffer from

difficulties in several aspects: strong parametric assumptions that lead to unrealistic synthesis results, high complexity in computation, and reliance on accurate age estimation (which is often not reliable). Recently, discriminative methods are proposed [14, 21, 22, 26]. The method in [22] uses gradient orientation pyramid (GOP) as feature and the support vector machine (SVM) as classifier for face recognition. In [21], SIFT [9] and Local Binary Pattern (LBP) [25] are used as features and a variation of random subspace LDA approach (RS-LDA) [38] is used for classification. Some variants of RS-LDA have also been used in [14, 26] for age invariant face recognition. These methods have been shown to be effective in certain cases. However, the lack of an underlying mechanism to capture facial structure across different ages may limit their generalizing performance.

In this paper, we consider a new approach to age-invariant face recognition. This approach is motivated by the belief that the facial image of a person can be expressed as combination of two components: an identity-specific component that is stable over the aging process, and the other component that reflects the aging effect. In particular, we introduce two latent factors: an *identity factor* and an *age factor*, which respectively govern the generation of these two components. Intuitively, each person is associated with a distinct identity factor, which is largely invariant over the aging process and thus can be used as a stable feature for face recognition; while the age factor changes as the person grows. For computational simplicity, we assume a linear model, where the identity components and the age components lie on two different subspaces. In this way, the problem of separating identity and age factors naturally reduces to a problem of learning the basis of these subspaces. As both the subspaces and the latent factors are unknown in the training stage, we derive an algorithm that can jointly estimate both from a set of training image, based on an Expectation-Maximization process. In this process, the latent factors and the model parameters are iteratively updated to maximize a unified objective. In the testing, given a pair of face images with unknown ages, we compute the match score between them by inferring and comparing the posterior mean of their identity factors.

The rest of the paper is organized as follows. Section 2 introduces the proposed HFA model. Section 3 presents the HFA-based age invariant face recognition framework. The experiments are presented in Section 4. Finally, Section 5 concludes this paper.

## 2. Hidden Factor Analysis

In this section, we propose a new model, called Hidden Factor Analysis (HFA), to address the problem of age invariant face recognition. We also develop an Expectation Maximization (EM) algorithm to estimate model parameters from data.

### 2.1. Problem Modeling

Matching facial images across ages is often necessary in real world applications. This is a challenging problem. As shown in Figure 1, faces of the same person can exhibit substantially different appearance at different ages, thus aggravating the difficulties. To address this problem, we consider a new approach, described below. This approach is primarily motivated by the observation that facial images (of different persons) usually share characteristics in common (e.g. skin and wrinkles). On the other hand, facial images of the same person also contain intrinsic features that are relatively stable across ages.

From a modeling standpoint, both age-specific and person-specific aspects can be respectively captured through latent factors, which we assume as statistically independent. Specifically, we use vectors to represent these latent factors, and call them *age factor* and *identity factor* throughout the remaining part of the paper. For simplicity and robustness, we consider a linear generative model, which expresses a facial image as a linear combination of three components: (1) age component, (2) identity component, and (3) a noise term which would allow actual observations to deviate from model space. In particular, the age component and identity component are respectively generated from the underlying age factor and identity factor through linear transformation. Overall, the joint model can be written as

$$\vec{t} = \vec{\beta} + U\vec{x} + V\vec{y} + \vec{\epsilon}. \quad (1)$$

We explain the terms of the model as follows.

1.  $\vec{t}$  is a  $d \times 1$  vector representing the observed facial features.
2.  $\vec{\beta}$  is a  $d \times 1$  vector representing the feature mean over the population.
3.  $\vec{x}$  is a  $p \times 1$  vector that represents the latent identity factor with prior distribution of  $N(0, I)$ .
4.  $\vec{y}$  is a  $q \times 1$  vector that represents the latent age factor with prior distribution of  $N(0, I)$ .
5.  $\vec{\epsilon}$  is a  $d \times 1$  vector that represents the additive noise, which follows an isotropic Gaussian where  $\vec{\epsilon} \sim N(0, \sigma^2 I)$ .
6.  $U$  is a  $d \times p$  matrix, whose columns span the subspace of cross-identity variation.
7.  $V$  is a  $d \times q$  matrix, whose columns span the subspace of cross-age variation.

The basic idea of our approach is to decompose facial features into identity components and age components based on this model, which are respectively generated from the identity factors and age factors. Below, we reiterate two properties of this model that are useful in our later discussion:

1. Any face feature  $\vec{t}$  consists of three components: the identity component  $U\vec{x}$ , age component  $V\vec{y}$ , and noise component  $\vec{\epsilon}$  representing noise and other variations in addition to age variations.
2. Through the decomposition based on this model, we can simultaneously attain two goals:  $U\vec{x}$  depends only on the subject's identity, with which we can perform age invariant face recognition, while  $V\vec{y}$  depends only on the subject's age, with which we can perform age estimation. However, pursuing this direction is beyond the scope of this paper. Instead, we focus on leveraging this decomposition to improve face recognition.

## 2.2. Model Adaptation

The model proposed in (1) has some parameters  $\theta = \{\vec{\beta}, U, V, \sigma^2\}$ . One can learn these parameters from data through maximum likelihood estimation, that is, to pursue a set of parameters that maximizes the following objective function:

$$Lc = \sum_{i,k} \ln p_{\theta}(\vec{t}_i^k, \vec{x}_i, \vec{y}_k). \quad (2)$$

$\vec{t}_i^k$  is the feature of the  $i$ -th subject at  $k$ -th age group.  $\vec{x}_i$  and  $\vec{y}_k$  are the corresponding identity and age factors, respectively. The summation is over all the available samples from different subjects at different age groups.

Note that this problem involves two latent variables  $\vec{x}_i$  and  $\vec{y}_k$ , which are not directly observable. Here, we adopt the coordinate ascent approach, alternately updating model parameters and latent factors with the other fixed. In particular, with the model parameter  $\theta$ , we can estimate the posterior distribution of the latent variables  $p_{\theta_0}(\vec{x}_i, \vec{y}_k|T)$ . Vice versa, we can update  $\theta$  given this distribution by maximizing the expectation of  $L$ . Given initial estimation  $\theta_0$ , our aim is to take a new estimation  $\theta$  to maximize:

$$\begin{aligned} \langle Lc \rangle = \\ \sum_{i,k} \int p_{\theta_0}(\vec{x}_i, \vec{y}_k|T) \ln p_{\theta}(\vec{t}_i^k, \vec{x}_i, \vec{y}_k) d\vec{x}_i d\vec{y}_k. \end{aligned} \quad (3)$$

Throughout the rest of this section, we present the detailed Expectation Maximization (EM) algorithm to iteratively update the model parameters.

### 2.2.1 Optimization for $\vec{\beta}$

According to equation (1),  $\vec{\beta}$  is the mean of the samples, which can be easily estimated by:

$$\vec{\beta} = \frac{1}{N} \sum_{i,k} \vec{t}_i^k, \quad (4)$$

where  $N$  is the total number of training samples.

### 2.2.2 Optimization for $U, V$ , and $\sigma^2$

To simplify the mathematical presentation, we will only give our conclusions in the following. The detailed proofs of them are attached in supplemental materials.

To optimize equation (3), we need to first obtain an estimation for joint distribution of the latent variables given model parameters  $\theta_0$ . The sufficient statistics of them are given as follows:

*Proposition 1\*<sup>1</sup>*: Given model parameter  $\theta_0$  and training data  $T = \{\vec{t}_i^k | i = 1, \dots, N; j = 1, \dots, M\}$ , the first and second conditional moments of  $p_{\theta_0}(\vec{x}_i|T)$ ,  $p_{\theta_0}(\vec{y}_k|T)$  and  $p_{\theta_0}(\vec{x}_i, \vec{y}_k|T)$  are given by:

$$\langle \vec{x}_i \rangle = \frac{U^T \Sigma^{-1}}{N_i} \sum_{k=1}^{N_i} (\vec{t}_i^k - \vec{\beta}) \quad (5)$$

$$\langle \vec{y}_k \rangle = \frac{V^T \Sigma^{-1}}{M_k} \sum_{i=1}^{M_k} (\vec{t}_i^k - \vec{\beta}) \quad (6)$$

$$\langle \vec{x}_i \vec{x}_i^T \rangle = \frac{I - U^T \Sigma^{-1} U}{N_i} + \langle \vec{x}_i \rangle \langle \vec{x}_i \rangle^T \quad (7)$$

$$\langle \vec{y}_k \vec{y}_k^T \rangle = \frac{I - V^T \Sigma^{-1} V}{M_k} + \langle \vec{y}_k \rangle \langle \vec{y}_k \rangle^T \quad (8)$$

$$\langle \vec{y}_k \vec{x}_i^T \rangle = -\frac{V^T \Sigma^{-1} U}{\sqrt{N_i M_k}} + \langle \vec{y}_k \rangle \langle \vec{x}_i \rangle^T \quad (9)$$

$$\langle \vec{x}_i \vec{y}_k^T \rangle = -\frac{U^T \Sigma^{-1} V}{\sqrt{N_i M_k}} + \langle \vec{x}_i \rangle \langle \vec{y}_k \rangle^T \quad (10)$$

where  $\Sigma = \sigma^2 I + U U^T + V V^T$ ,  $N_i$  and  $M_k$  are the numbers of training samples for the  $i$ -th subject and the  $k$ -th age group, respectively (e.g. if we have 100 training samples fall into the  $k$ -th age group, then  $M_k$  is 100).

In the Expectation step, we compute the sufficient statistics according to *Propositional 1*, and in the Maximization step, we update the model parameters to maximize equation (3), as described in the following:

*Proposition 2\*<sup>2</sup>*: Given initial model parameter  $\theta_0$ , the new parameter  $\theta$  maximizes the expectation of the log-likelihood defined in equation (3) if:

$$U = (C - DB^{-1}E)(A - FB^{-1}E)^{-1} \quad (11)$$

$$V = (D - CA^{-1}F)(B - EA^{-1}F)^{-1} \quad (12)$$

$$\sigma^2 = \frac{1}{Nd} \sum_{i,k} \left\{ (\vec{t}_i^k - \vec{\beta} - U \langle \vec{x}_i \rangle - V \langle \vec{y}_k \rangle)^T (\vec{t}_i^k - \vec{\beta}) \right\} \quad (13)$$

<sup>1</sup>Proof for Proposition 1 is attached in supplemental materials

<sup>2</sup>Proof for Proposition 2 is attached in supplemental materials

where

$$A = \sum_{i,k} \langle \vec{x}_i \vec{x}_i^T \rangle \quad B = \sum_{i,k} \langle \vec{y}_k \vec{y}_k^T \rangle$$

$$C = \sum_{i,k} \left( \vec{t}_i^k - \vec{\beta} \right) \langle \vec{x}_i \rangle^T \quad D = \sum_{i,k} \left( \vec{t}_i^k - \vec{\beta} \right) \langle \vec{y}_k \rangle^T$$

$$E = \sum_{i,k} \langle \vec{y}_k \vec{x}_i^T \rangle \quad F = \sum_{i,k} \langle \vec{x}_i \vec{y}_k^T \rangle$$

$N$  is the total number of training samples, and  $d$  is the dimension of  $\vec{\beta}$  as above.

Equations (11)-(13) indicate that the latent variables enter into the EM formulation only through sufficient statistics, as expected. The EM algorithm jointly estimates the distribution of both the identity and age components (which are believed to be associated to both the identity factors and age factors controlling the generation of facial appearance), and then adapts the model parameters such that the expectation of the log-likelihood that these training samples are generated from these latent factors through the model is maximized. The convergence proof for this algorithm is attached in the supplemental materials. Algorithm 1 summarizes the procedure of the EM algorithm for the HFA model adaptation. The computationally intensive part lies in step 2) and 3). Each iteration involves one matrix inversion of size  $d \times d$ , two inversions of  $p \times p$  and  $q \times q$  (see Table 3 for detailed settings). The algorithm converges quickly within 10 iteration steps in our experiments.

---

**Algorithm 1** EM algorithm for Hidden Factor Analysis.

---

**Input:** feature set of training images  $D = \left\{ \vec{t}_i^k \right\}$ , with identity labels and age group information.

**Output:** model parameter  $\theta = \left\{ \vec{\beta}, U, V, \sigma^2 \right\}$

**Initialization:**

$$\sigma^2 \leftarrow 0.1$$

$$U \leftarrow \text{rand}(-0.1, 0.1)$$

$$V \leftarrow \text{rand}(-0.1, 0.1)$$

- 1) Compute  $\vec{\beta}$  with equation (4).
  - 2) Compute the sufficient statistics with the equations (5)-(10).
  - 3) Update the model parameters with the equations (11)-(13).
  - 4) Go to step (2) until convergence.
- 

### 2.3. Discussion

1. Our approach differs significantly from the prior subspace factoring works [16] [23] [4]. The supervised approach in learning the optimal factor spaces distinguishes our work from [16] [23] [4] that mostly use unsupervised

way in factor analysis.

2. Our approach can generalize to other face recognition scenarios. The proposed model in (1) decomposes the original face feature into three components: the common feature component, the variation component, and the noise component. By retrieving the common component, we can reduce the variations contained within subspace  $V$  as well as noise component. For general applications, such as matching faces in the wild, we can replace the aging variations with other kinds of variations.

## 3. HFA based Age invariant Face Recognition Framework

In this section, we present our age invariant face recognition framework based on the proposed HFA model in Section 2. The organization of this framework is illustrated in Figure 2. We describe each framework component in the following subsections.

### 3.1. Local Feature Representation

Local facial features have been shown to be more effective than the global facial features in representing face images at various scales and orientations. Among all the existing local feature descriptors, Histograms of Oriented Gradients (HOG) [6] is one of the most successful ones. Thus, in our experiment we will apply HOG as the feature descriptor. For any face image, we first divide it into a set of overlapping patches, and then apply the HOG descriptor on each patch to extract the HOG features. The extracted HOG features from all the patches are concatenated together to form a long feature vector for further analysis. The details of our setting are listed in Table 1.

Table 1. Parameter settings for local feature representation.

Overlapping Factor	Patch Size	#Orientations	Scales <sup>1</sup>
0.5	12	12	1.0, 1.5, 2.0

Prior to applying the HOG feature extractor, we preprocess all the face images through the following steps:

- 1) Rotate the face images to align them to the vertical orientation;
- 2) Scale the face images so that the distances between the two eyes are the same for all images;
- 3) Crop the face images to remove the background and hair region;
- 4) Apply histogram equalization to the cropped face images for photometric normalization.

---

<sup>1</sup>Scales: sampling distance in pixel, with bilinear interpolation.



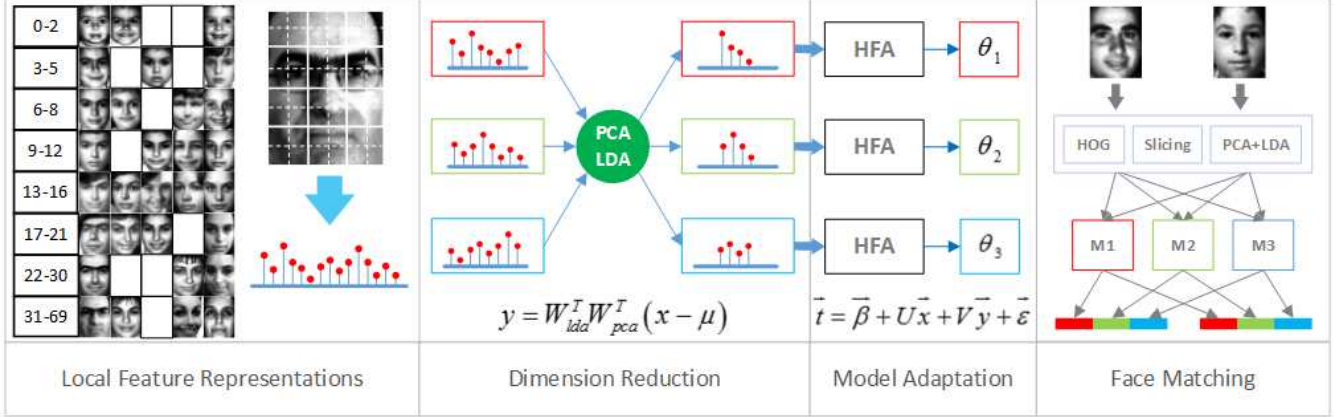


Figure 2. Illustration of the HFA based age-invariant face recognition system. At the training stage, the training faces are first grouped according to their *identities* and *ages* (corresponding to index  $i$  and  $k$  in Algorithm 1, respectively), followed by feature extraction (section 3.1) on each image. With each training face represented by HOG feature, we reduce the dimension of these features with slicing (three slices are shown in the figure), PCA and LDA (section 3.2). Finally, HFA models are adapted independently on each of the sliced features of the dataset (section 2.2), obtaining a set of model parameters for each slice. At the testing stage, the matching score of the given face pair (one from probe and the other one from gallery) is computed by first going through feature extraction and dimension reduction steps the same as training, then estimating the identity latent variables for each slice of the two face features. The final matching score is given by the cosine distance of the concatenated identity features (section 3.3).

### 3.2. Dimension Reduction

Due to the use of the multi-scale and densely sampling techniques in local feature representation stage, the extracted feature vector is of very high dimension (e.g. up to  $10^6$  in our experiment). Such a high dimensionality presents huge demands on computer memory and significantly slows down the processing speed. It is desirable to perform dimension reduction prior to applying our HFA model. We use two simple yet effective techniques for this purpose: feature slicing and subspace analysis. Inspired by the multi-level subspace analysis techniques in [32] [33] [34] [20] [15], we first divide the long feature vector into several slices equally, and then apply PCA [36] + LDA [5] [39] (which is a widely used dimension reduction technique in face recognition) on each slice to obtain a compressed slice of smaller feature vector for subsequent analysis.

### 3.3. Face Matching

After local feature representation and dimension reduction, for each face image we have several compressed slices of smaller feature vectors. Based on each slice, we construct a HFA model. In the matching process, for each pair of probe sample and gallery sample, we first compute the predictive distribution of their identity variables, as follows:

$$P(\vec{x}|\vec{t}) = \int P(\vec{x}|\vec{t}, \vec{y}) P(\vec{y}|\vec{t}) d\vec{y}.$$

The predictive distribution is given by:

$$P(\vec{x}|\vec{t}) = N(\vec{x}|U^T \Sigma^{-1}(\vec{t} - \vec{\beta}), I - U^T \Sigma^{-1} U),$$

where  $\Sigma = \sigma^2 I + U U^T + V V^T$  as before. Then, we use the  $U \langle \vec{x}|\vec{t} \rangle$  with cosine distance for face matching. Note that we concatenate the output of each HFA model to form the final feature for classification. The matching process is summarized as follows:

1. Apply the HFA model on each slice of data, and obtain the output feature vector:  $\vec{f}_i = U U^T \Sigma^{-1}(\vec{t} - \vec{\beta})$  from the  $i$ -th model.
2. Concatenate the feature vectors to form a long vector:  $\vec{F} = [\vec{f}_1 \dots \vec{f}_K]$ .
3. Compute the matching score using the cosine distance.

Figure 2 illustrate the entire pipeline of our system. The matching process of our HFA model does not need any age information of the test images. So the HFA model can be easily used in practical applications.

## 4. Experiments

### 4.1. Database

There are two well-known public domain databases for age invariant face recognition: MORPH [13] and FGNET [2]. The MORPH database has two separate datasets: Album 1 and Album 2. The MORPH Album 1 only contains 1690 face images from 625 different subjects. The MORPH Album 2 is the largest face aging dataset available in the public domain. This dataset is composed of about 78,000 face images of 20,000 different subjects captured at different ages. Comparing to the MORPH Album 1 dataset, the MORPH Album 2 dataset has two desired attributes: (i)

very large number of subjects, and (ii) large number of face images captured at different ages. So we use the MORPH Album 2 dataset in this study. The FG-NET dataset consist of 82 different individuals, with each one having multiple images (13 on average) taken at different age levels. Figure 3 shows the age range distribution for these two datasets. In this paper, we will use both the MORPH Album 2 and FGENT for our experimental validation.

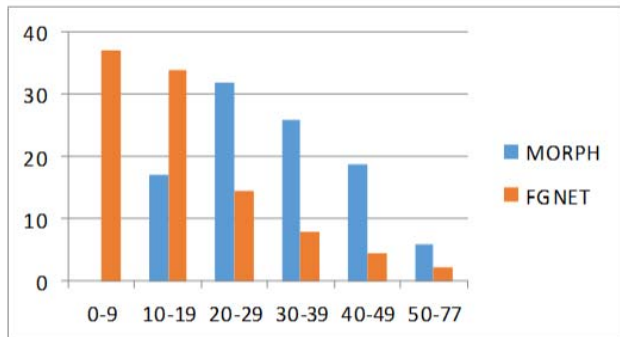


Figure 3. Age range distribution (%) of MORPH Album 2 and FG-Net.

To train our HFA model, we first partition the training data set into several age groups. To balance the number of training samples in each age group, we partition the age into 8 groups such that each group has approximately the same number of samples, as shown in Table 2.

Table 2. Age groups partition scheme. Ages are divided into eight groups, with #samples shown.

Groups	Age range		#Samples	
	MORPH	FG-NET	MORPH	FG-NET
1	15-19	0-2	5390	109
2	20-23	3-5	5829	124
3	24-26	6-8	4106	113
4	27-31	9-12	5343	135
5	32-36	13-16	5395	131
6	37-40	17-21	4013	134
7	41-46	22-30	5292	127
8	47-77	31-69	4632	129

## 4.2. Parameter Settings

The HFA model has some free parameters:  $d$  (the dimension of the feature vector fed into the model),  $p$  (the dimension the identity factor), and  $q$  (the dimension of the age factor). In addition to these parameters, there are some other parameters for our systems: number of slices (see section 3.2), dimension of PCA [36] and LDA [5] [39] (see section 3.2), as well as the size of the normalized face images (see section 3.1). The detailed parameter settings on the datasets are shown in Table 3.

Table 3. The parameters used in our experiments.

Parameters		MORPH	FG-NET
HFA Model	$d$	1000	150
	$p$	100	80
	$q$	15	10
System	#Slice	6	3
	PCA dim	1001	150
	LDA dim	1000	N/A
	Image size	$200 \times 150$	$100 \times 75$

## 4.3. Experiment on the MORPH Ablum 2 dataset

The MORPH Album 2 dataset is the largest publicly available face aging dataset. It is very crucial to conduct large-scale experimental validation to evaluate our HFA model on this dataset. Following the configurations of training and testing split in [21], we use all the 20,000 persons for this experiment. We partition the MORPH album 2 dataset into a training set and an independent test set. The training data consists of 20,000 face images from 10,000 subjects, with each subject having two images with the largest age gap. The test data is composed of a gallery set and a probe set from the remaining 10,000 subjects. The gallery set is composed of 10,000 face images corresponding to the youngest age of these 10,000 subjects. The probe set is composed of 10,000 face images corresponding to the oldest age of these 10,000 subjects.

We compare our HFA model against several state-of-the-art methods for age invariant face recognition on MORPH Album 2. They include (i) FaceVACS, a leading commercial face recognition engine [5], (ii) several newly developed generative methods [7, 27] for face aging, and (iii) several newly developed discriminative methods [14, 21, 26] for direct age invariant face recognition. The comparative results are reported in Table 4. All the methods in Table 4 are tuned to the best settings according to their papers. From these results, we have the following observations. First, among the published methods in Table 4, the MFDA method in [21] is the best-performing one. This shows the effectiveness of the discriminative model. It is encouraging to see that our approach significantly outperforms the MFDA method by improving the rank-1 identification rate from 83.90% to 91.14%. To our best knowledge, this is the best identification rank-1 result on such a large-scale matching scenario (using 10,000 face images as the gallery set and another 10,000 face images as the probe set from 10,000 different persons) in the MORPH Album 2 dataset.

Finally, we show some examples of failed retrievals in Figure 4. While the rank-1 retrievals are not correct in these cases, the probe images appear to be more similar to the incorrect rank-1 matched images than the true images.

Table 4. Rank-1 identification rates on the MORPH Album 2 dataset.

Algorithms	Recognition Rates
FaceVACS [1]	78.90%
Park et al. (2010) [27]	79.80%
Du et al. (2012) [7]	79.24%
Li et al. (2011) [21]	83.90%
Klare et. al. (2011) [14]	79.08%
Otto et al. (2012) [26]	81.27%
<b>HFA</b>	<b>91.14%</b>

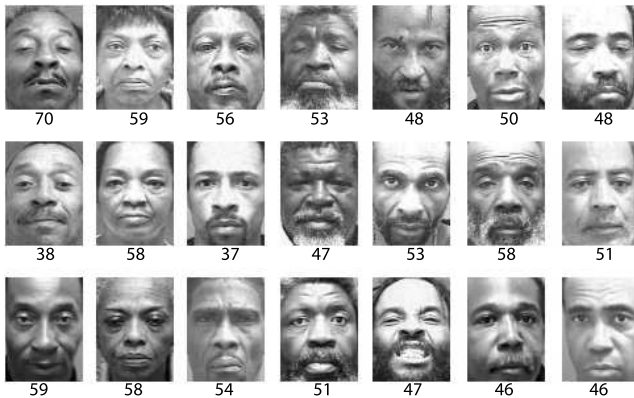


Figure 4. Some examples of failed retrievals in MORPH Album 2. The first row presents the probe faces, the second row is the incorrect rank-1 matching results using our approach, and the bottom row shows the corresponding ground-truth faces for the probes.

#### 4.4. Experiment on the FGNET dataset

The good performance of our HFA model in Table 4 is particularly interesting. It is desirable to investigate the generalization ability of our HFA model across different datasets. To this end, we conduct an additional experiment to compare our HFA model with the state-of-the-art results on the FGNET dataset. Following the training and testing split scheme in [21], we use the leave-one-person-out scheme for experimental validation. The comparative results are reported in Table 5, from which we can also see the significant performance improvement of our HFA model over the state-of-the-art. This confirms the effectiveness of our algorithm.

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

Algorithms	Recognition Rates
FaceVACS [1]	26.4%
Park et al. (2010) [27]	37.4%
Li et al. (2011) [21]	47.5%
<b>HFA</b>	<b>69.0%</b>

Figure 5 shows some failed retrievals in the FGNET dataset. These results confirm what we observed from Fig-

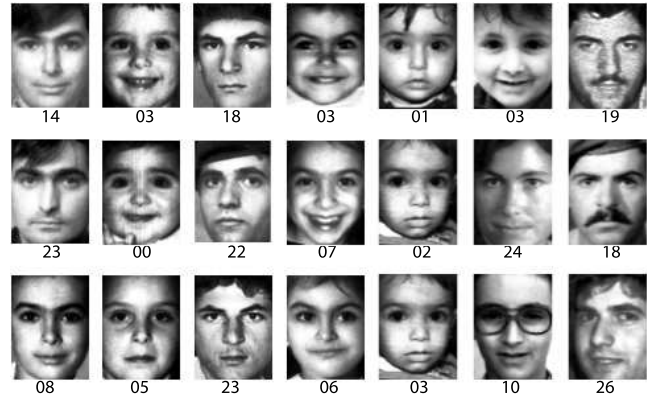


Figure 5. Some examples of failed retrievals in FGNET. The first row presents the probe faces, the second row is the incorrect rank-1 matching results using our approach, and the bottom row shows the corresponding ground-truth faces for the probes.

ure 4: The rank-1 retrieved images appeared highly similar to the probe images in the incorrect matchings.

## 5. Conclusion

In this paper, we have proposed a hidden factor analysis (HFA) approach to address the challenging problem of age invariant face recognition. The basic idea of the HFA model is to separate the aging variations from the person-specific features for pursuing the robust age-invariant face features. Extensive experiments conducted on two public domain face aging datasets (MORPH Album 2 and FGNET) convincingly demonstrate the superiority of our HFA model over the state-of-the-art algorithms.

## 6. Acknowledgements

This work was supported by grants from Natural Science Foundation of China (61103164, 61070148), Guangdong Innovative Research Team Program (No. 201001D0104648280), and the General Research Fund sponsored by the Research Grants Council of the Kong Kong SAR (Project No. CUHK 416312).

## References

- [1] Facevacs software developer kit, cognitec systems GbmH. 7
- [2] FG-NET Aging Database. 1, 5
- [3] A. Albert, K. Ricanek Jr, and E. Patterson. A review of the literature on the aging adult skull and face: implications for forensic science research and applications. *Forensic Sci Int*, 172(1):1–9, 2007. 1
- [4] A. C. B. J. Frey and T. S. Huang. Mixtures of local linear subspaces for face recognition. In *CVPR*, 1998. 4
- [5] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 19(7):711–720, 1997. 5, 6



- [6] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *CVPR (1)*, pages 886–893, 2005. 4
- [7] J.-X. Du, C.-M. Zhai, and Y.-Q. Ye. Face aging simulation based on NMF algorithm with sparseness constraints. *Neurocomputing*, 2012. 1, 6, 7
- [8] Y. Fu and T. S. Huang. Human age estimation with regression on discriminative aging manifold. *IEEE Transactions on Multimedia*, 10(4):578–584, 2008. 1
- [9] C. Geng and X. Jiang. Face recognition using sift features. In *ICIP*, pages 3313–3316, 2009. 2
- [10] X. Geng, Z.-H. Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(12):2234–2240, 2007. 1
- [11] G. Guo, Y. Fu, C. R. Dyer, and T. S. Huang. Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Transactions on Image Processing*, 17(7):1178–1188, 2008. 1
- [12] G. Guo, G. Mu, Y. Fu, and T. S. Huang. Human age estimation using bio-inspired features. In *CVPR*, pages 112–119, 2009. 1
- [13] K. R. Jr. and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. In *FG*, pages 341–345, 2006. 5
- [14] B. Klare and A. K. Jain. Face recognition across time lapse: On learning feature subspaces. In *IJCB*, pages 1–8, 2011. 2, 6, 7
- [15] B. Klare, Z. Li, and A. K. Jain. Matching forensic sketches to mug shot photos. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(3):639–646, 2011. 5
- [16] J. H. Kuang-Chih Lee and D. J. Kriegman. Acquiring linear subspaces for face recognition under variable lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005. 4
- [17] Y. H. Kwon and N. D. V. Lobo. Age classification from facial images. In *In Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pages 762–767, 1999. 1
- [18] A. Lanitis, C. Draganova, and C. Christodoulou. Comparing different classifiers for automatic age estimation. *IEEE Trans Syst Man Cybern B Cybern*, 34(1):621–8, 2004. 1
- [19] A. Lanitis, C. J. Taylor, and T. F. Cootes. Toward automatic simulation of aging effects on face images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(4):442–455, 2002. 1
- [20] Z. Li, D. Lin, and X. Tang. Nonparametric discriminant analysis for face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(4):755–761, 2009. 5
- [21] Z. Li, U. Park, and A. K. Jain. A discriminative model for age invariant face recognition. *IEEE Transactions on Information Forensics and Security*, 6(3-2):1028–1037, 2011. 2, 6, 7
- [22] H. Ling, S. Soatto, N. Ramanathan, and D. W. Jacobs. Face verification across age progression using discriminative methods. *IEEE Transactions on Information Forensics and Security*, 5(1):82–91, 2010. 2
- [23] N. A. M. H. Yang and D. Kriegman. Face detection using mixtures of linear subspaces. In *FG*, 2000. 4
- [24] A. Montillo and H. Ling. Age regression from faces using random forests. In *ICIP*, pages 2465–2468, 2009. 1
- [25] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987, 2002. 2
- [26] C. Otto, H. Han, and A. K. Jain. How does aging affect facial components? In *ECCV Workshops (2)*, pages 189–198, 2012. 2, 6, 7
- [27] U. Park, Y. Tong, and A. K. Jain. Age-invariant face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(5):947–954, 2010. 1, 6, 7
- [28] N. Ramanathan and R. Chellappa. Face verification across age progression. *IEEE Transactions on Image Processing*, 15(11):3349–3361, 2006. 1
- [29] N. Ramanathan, R. Chellappa, and S. Biswas. Computational methods for modeling facial aging: A survey. *J. Vis. Lang. Comput.*, 20(3):131–144, 2009. 1
- [30] J.-L. Suo, X. Chen, S. Shan, and W. Gao. Learning long term face aging patterns from partially dense aging databases. In *ICCV*, pages 622–629, 2009. 1
- [31] J.-L. Suo, S. C. Zhu, S. Shan, and X. Chen. A compositional and dynamic model for face aging. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(3):385–401, 2010. 1
- [32] X. Tang. Texture information in run-length matrices. *IEEE Trans. Image Processing*, 1998. 5
- [33] X. Tang and Z. Li. Frame synchronization and multi-level subspace analysis for video based face recognition. In *CVPR*, 2004. 5
- [34] X. Tang and Z. Li. Audio-guided video based face recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 2009. 5
- [35] N. Tsumura, N. Ojima, K. Sato, M. Shiraishi, H. Shimizu, H. Nabeshima, S. Akazaki, K. Hori, and Y. Miyake. Image-based skin color and texture analysis/synthesis by extracting hemoglobin and melanin information in the skin. *ACM Trans. Graph.*, 22(3):770–779, 2003. 1
- [36] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*, pages 586–591, 1991. 5, 6
- [37] J. Wang, Y. Shang, G. Su, and X. Lin. Age simulation for face recognition. In *ICPR (3)*, pages 913–916, 2006. 1
- [38] X. Wang and X. Tang. Random sampling lda for face recognition. In *CVPR (2)*, pages 259–265, 2004. 2
- [39] X. Wang and X. Tang. A unified framework for subspace face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(9):1222–1228, 2004. 5, 6
- [40] S. Yan, H. Wang, X. Tang, and T. S. Huang. Learning auto-structured regressor from uncertain nonnegative labels. In *ICCV*, pages 1–8, 2007. 1
- [41] S. K. Zhou, B. Georgescu, X. S. Zhou, and D. Comaniciu. Image based regression using boosting method. In *ICCV*, pages 541–548, 2005. 1