Discriminative Aging Subspace Learning for Age Estimation

Manisha Sawant (mpparlewar@gmail.com)
VNIT: Visvesvaraya National Institute of Technology https://orcid.org/0000-0002-3682-4146
Kishor Bhurchandi
Visvesvaraya National Institute of Technology

Research Article

Keywords: Age estimation, Hidden factor analysis (HFA), Aging manifold, Regression

Posted Date: August 4th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-679477/v1

Noname manuscript No.
(will be inserted by the editor)

Discriminative Aging Subspace Learning for Age Estimation

Manisha M. Sawant and Kishor M. Bhurchandi

Received: date / Accepted: date

Abstract Hidden factor analysis (HFA) has been widely used in age invariant face recognition systems. It decomposes facial features into independent age factor and identity factor. Age invariant face recognition systems utilize identity factor for face recognition; however, the age factor remains unutilized. The age component of the hidden factor analysis model depends on the subject's age, hence it carries a significant age related information. In this paper, we propose the HFA model based discriminative manifold learning method for age estimation. Further, multiple regression methods are applied on low dimensional features learned from the aging subspace. Extensive experiments are performed on a large scale aging database MORPH II and the accuracy of our method is found superior to the current state-of-the-art methods.

Keywords Age estimation · Hidden factor analysis (HFA) · Aging manifold · Regression

1 Introduction

Human face conveys significant information for human-machine interaction. Estimating various facial attributes such as age, gender and expression plays a vital role in various forensic, multimedia and law enforcement applications. Among various attributes such as gender, age and expression, age plays an important role in social interations and health care systems. Children and senior citizens require special attention and care in social interactions. Various organizations such as immigration department, driving licence office and forensic department requires knowledge of the applicant's age. Due to the various applications in health care, law enfircement and governance, analysis of facial aging has gained major attention in computer vision and machine learning community. Facial aging related research is broadly classified into three categories: age estimation, age progression and cross age face recognition. Age estimation and progression mainly focuses on aging information that changes due to aging. Whereas, in the cross age face recognition major emphasis is on identity information that does not change with aging. Recent works on face recognition [12, 13, 24, 34, 35, 42, 45, 48] adopt hidden factor analysis [12] and probabilistic linear discriminant analysis (PLDA) [34] to separate within subject and between subject information from facial images. These methods build a probabilistic linear model and iteratively learn optimal model parameters using expectation-maximization (EM) [31] algorithm. Such decomposition of facial features has also been used to separate age and identity components to achieve cross-age face recognition [12, 13, 24, 45, 48] and age progression [28, 42, 48]. Aging face recognition utilizes the identity factor, whereas, both age and identity factors are used for rendering age progressed images. It is clear from the performance of these methods that identity factor provides identity information that is stable across age progression. However, the information conveyed by the aging component remains unutilized; this work utilizes the same.

Apart from significant research on cross-age face recognition [12, 13, 24–26, 32, 45], relatively few publica-

M. Sawant (M.S.) (corresponding author) · K.M. Bhurchandi (K.M.B.)
Visvesvaraya National Institute of Technology, Nagpur, India

E-mail: mpparlewar@gmail.com

bhurchandikm@ece.vnit.ac.in

tions have been reported on age estimation [5,8,11,17, 33,43,49]. This is due to various factors such as facial aging is a complex biological change that affects the shape and texture of the face. Different aging patterns are observed (even in the people of same ethnic origins) due to diversity in climatic conditions and lifestyle. Hence, it is difficult to precisely predict a persons age from the facial appearances, even for humans.

Key components in automatic age estimation methods are facial feature representation and age prediction. Performance of the age estimation system varies with the quality of the facial feature representation. In literature, we find various local feature descriptors that are used for extracting appearance features from an image. Guo et al. [17] proposed Bio-Inspired Feature (BIF) to extract age related facial features for age estimation. BIF and its variants have been used in [15, 16,19] for age estimation. Various existing histogram based local features have also been widely used in age estimation [9,33,46,50]. Pontes et al. combined local and global features in [33], where Active Appearance Model (AAM) [6] is used as a global feature and Local Binary Pattern (LBP)[1], Gabor and Local Phase Quantization (LPQ) [2] are used as local features. Histogram of Oriented Gradients (HOG) is used in [20, 50] for age estimation. Scale Invariant Feature Transform (SIFT) [29] and LBP are used as feature vectors in [47] for global age estimation approach. Apart from existing local feature descriptors, a few local feature descriptors specially designed for age estimation include Directional Age-Primitive Pattern (DAPP) [22] and Local Direction and Moment Pattern (LDMP) [38].

Age estimation has been considered either as a classification problem or a regression problem. In classification, age label is considered as a class, whereas, in regression it is treated as a continuous value. While classification tries to learn from age variations in the training data, regression utilizes the label information. Recently proposed learning algorithms for age estimation include support vector machine (SVM)/regression (SVR) [19], Partial Least Squares (PLS) [16] along with its kernelized version (KLPS) [15] and Canonical Correlation Analysis (CCA) [16]. Ordinal hyperplane ranking algorithm (OHRank) [5] utilizes the relative order information among age labels. Apart from above mentioned regression methods, multi-task Warped Gaussian Process (WGP) regression has been proposed [49] for personalized age estimation. Computationally efficient extension of WGP has been proposed in [50] using Orthogonal Gaussian Process (OGP). To further improve computation complexity and age estimation performance, a two level hierarchical Gaussian process regression was proposed in [39].

As discussed earlier in this section, hidden factor analysis is a probabilistic model which separates identity and age components from the facial features. The HFA and its modified versions are widely used in many state-of-the-art age invariant face recognition systems [12,24,45]. All these methods use identity factor and ignore the information present in the second hidden factor *i.e.* age factor. The age component contains important information about the subjects age. Therefore, the age component of the HFA model is hypothesized to accelerate the age estimation performance. Also the age component on the feature space is low dimensional. Therefore, to estimate age from the face images, we propose a novel manifold learning method derived from the HFA model. We learn the low dimensional aging features from the learned aging subspace. The aging features learned from the aging subspace carries age discriminative information. Thus the regression algorithm can fit the data better and is expected to improve the age estimation performance.

In this paper, we propose a novel scheme for aging feature extraction from discriminative aging manifold for automatic age estimation. The basic idea is to learn a low-dimensional embedding of the aging manifold using the HFA based subspace learning method. Further we use multiple regression techniques for subsequent age estimation. The efficiency of the proposed method has been demonstrated with extensive experiments on multiple regression algorithms using a large scale age database. The contributions of this paper are given as follows. First, our method is the first to present the HFA based discriminative aging subspace. Second, we empirically show that our method provides a robust discriminative aging subspace, which provides consistent results on multiple regression methods.

This paper is organized as follows. We provide detailed related work in section 2. Proposed age estimation framework and manifold learning is presented in section 3. The experimental setup and results are discussed in section 4. We conclude the paper in section 5.

2 Aging Manifold

Manifold based feature extraction has gained a considerable interest over a past few years. The basic idea of manifold features extraction is that high-dimensional data lie on or close to a smooth low-dimensional manifold. Some representative works on manifold learning include Locally Linear Embedding (LLE) [37], ISOMAP [40], and Laplacian eigenmaps [3]. However, it has been observed that, these methods can only work well on the training samples and not on test data.

In age manifold, a common aging pattern is learned from images of many individuals with different ages. Several face images are adopted to represent an age. Each subject may be represented by one image or several images at different ages. These images make a set referred to as a manifold which makes up points in a high dimensional vector space. Several facial images are utilized to learn a non-linear low-dimensional aging pattern. It is assumed that the faces with close ages are located closely on the manifold, as the aging face images are distributed on to an intrinsic low-dimensional manifold. Individuals may have as low as one image at each age in a database. Once the low dimensional aging manifold is learn, a regression is applied on the embedded subspace to predict exact age.

In the recent years, many efforts have been devoted to identify discriminant aging subspaces for age estimation. A subspace called AGing pattErn Subspace (AGES) is proposed by Geng et al. [9,10] to learn personalized aging process from multiple faces of an individual. Discriminative subspace learning based on aging manifold and Locally Adjusted Robust Regressor (LARR) was proposed by, Guo et al. [14]. In LARR, the ages are first predicted by a regression function. Then the predicted age value is fine tuned locally to match with the true value within a specific bound. This method shows that the facial images can be represented on a manifold. The age manifold method does not require images at different ages of the same individual, like the AGES method. But in order to learn the low-dimensional manifold, it requires images of several subjects at many ages. Later, Kernel Partial Least Square Regression (KPLS) [15] has been used that simultaneously reduces the dimension and learns the aging function. The performance of these methods depend on how well the manifold is learnt from the facial images to precisely represent the age. Thus identifying a space where aging manifold truly lies is still an open research problem. The age manifold also requires large database in order to satisfactorily learn the aging information. Existing manifold based methods extract manifold features from the gray intensity or image space. However, the image space is inefficient to model the large age variations. The texture features such as LBP, HOG, SIFT and LPQ are used to capture the textural variations due to aging. But the manifold of such feature space has not been explored for age estimation so far. In this paper, we propose a subspace learning scheme for aging feature extraction for automatic age estimation. The basic idea is to learn an aging manifold embedding using the HFA based aging subspace learning method.

3 Proposed Work

The traditional age estimation approaches basically use generalised local feature descriptors used in facial analysis [9,33,46,50], or design the age-related characteristics by extracting edge and directional information from the training samples [17,22,38], without considering age group specific aging patterns. A few methods design person specific age estimation schemes [10,49], which assume that every person ages in a different way. Such assumption makes accuracy of the system, identity dependent and poses restriction on the age estimation system. These facts and an objective of discriminative aging feature motivate us to design an approach to accurately learn low dimensional identity independent aging features. The proposed method, learns low dimensional aging features by separating person-specific component. It necessarily separates identity information and the age-related component that conveys the age from the local feature descriptor. In the proposed method only the latter component plays an important role.

In this section, the framework of the proposed manifold learning is introduced. Given a set of local features of face images from diverse age groups, for manifold learning, a hidden factor analysis method, is introduced in this work. Training set images are used to learn the bases of the age and identity subspace. Once the parameters are well trained, by projecting the training and test faces onto the aging subspace, the age-discriminative features can be achieved for every face image. It is well known that, histogram based local feature representation often leads to high dimensionality which may result into the failure of regressors. Therefore, an appropriate manifold learning algorithm is required to reduce redundancies in the feature dimension. In this paper, we propose a discriminative aging subspace learning procedure using hidden factor analysis.

3.1 Aging Subspace Learning based on HFA

Hidden factor analysis [12] is a linear probabilistic model that separates identity specific variations from the age for age invariant face recognition. The probabilistic linear factor analysis model presented in [12] is as follows:

$$f = \mu + \mathbf{Q}x + \mathbf{P}y + \varepsilon \tag{1}$$

The major notations used are defined as follows:

- Observable feature vector f is denoted by a $D \times 1$ vector, where D is dimension of feature vector.
- Mean of the face features μ is denoted by a $D \times 1$ vector.

- Hidden identity factor x is a $r \times 1$ random vector. A prior distribution of identity factor is defined as, $p(x) \sim \mathcal{N}(0, I_r)$, where $\mathcal{N}(\mu, K)$ denotes multivariate Gaussian distribution with mean μ and covariance matrix K.
- Hidden age factor y is a $s \times 1$ random vector. A prior distribution of age factor is defined as, $p(y) \sim \mathcal{N}(0, I_s)$.
- Cross-identity variations are captured by a $D \times r$ identity factor loading matrix $\mathbf{Q}$.
- Cross-age variations are captured by a $D \times s$ age factor loading matrix $\mathbf{P}$.
- Observation noise denoted by ε follows an isotropic Gaussian distribution i.e. $\varepsilon \sim \mathcal{N}\left(0, \sigma^2 I\right)$.

The HFA model in (1) decomposes the observed feature into identity, age and noise components. EM algorithm is used to estimate the model parameters $\left\{\mu, \mathbf{Q}, \mathbf{P}, \sigma^2\right\}$. Such a decomposition model has been effectively used for age invariant face recognition [12,13,24,45,48].

We decompose facial features f_i^k into linear combination of identity, age and noise component using HFA model. Suppose that f_i^k represents the feature vector of the individual i at age-group k, then the discriminative aging subspace is learned from (2).

$$f_i^k = \mu + \mathbf{Q}x_i + \mathbf{P}y_k + \varepsilon \tag{2}$$

Since the observed feature space does not provide efficient representation for age estimation, we offer a robust representation with fine properties. The second component of HFA model in (2) depends only on the subjects identity, whereas the third component depends only on the age of the subject. Matrix $\mathbf{P}$ in (2) is age factor loading matrix that captures the cross-age variations. The columns of the matrix $\mathbf{P}$ are the bases for cross-age variations; therefore we term matrix $\mathbf{P}$ as aging subspace. The age factor y_k can be viewed as the position of the f_i^k *i.e.* observation variable in the aging subspace. The relationship between the observation space f_i^k and the identity and aging subspace is indicated in Fig. 1.

The learning objective is to estimate the model parameter $\theta = \left\{\mu, \mathbf{Q}, \mathbf{P}, \sigma^2\right\}$ using training data. We learn the set of parameters by maximizing joint likelihood $p_\theta\left(f_i^k, x_i, y_k\right)$ of the observed feature vector and associated identity and age factors. Both the identity and age factors are not observed directly. It is possible to only infer them through their posterior distributions for fixed set of model parameters. EM is used to solve the optimization problem. EM algorithm is a maximum likelihood method that iteratively updates set of parameters. The objective function for optimization is,

$$L_c = \sum_{i,k} ln\left(p_\theta\left(f_i^k, x_i, y_k\right)\right). \tag{3}$$

Given the initial estimate of parameters θ_0, in Expectation-step (i.e. E-step) first two moments of the age factor are computed as

$$\mu = \frac{1}{N}\sum_{i,k} f_i^k \tag{4}$$

$$E(y_k) = \frac{\mathbf{P}^T \sum^{-1}}{N_k}\sum_{i=1}^{N_k}\left(f_i^k - \mu\right) \tag{5}$$

$$E\left(y_k y_k^T\right) = \frac{I - \mathbf{P}^T \sum^{-1}\mathbf{P}}{N_k} + E(y_k)\,E(y_k)^T \tag{6}$$

where $\sum = \mathbf{Q}\mathbf{Q}^T + \mathbf{P}\mathbf{P}^T + \sigma^2 I$, N is total training images, N_k is number of images at k^{th} age group, and $E[\cdot]$ is expectation. The estimate of the first two moments of the identity factor is calculated in a similar way. In Maximization-step (i.e. M-step), the model parameters are computed using the estimate of age and identity factors. The complete log likelihood in (3) is rewritten as,

$$L_c = \sum_{i,k} ln\left(p\left(f_i^k | x_i, y_k\right)\right) + ln\left(p\left(x_i, y_k\right)\right)$$

where $p\left(f_i^k | x_i, y_k\right) = \mathcal{N}\left(\mu + \mathbf{Q}x_i + \mathbf{P}y_k, \sigma^2\right)$ and $p(x_i, y_k) = \mathcal{N}(0, I)$.

In M-step, first the derivative of the objective function is computed and the optimal parameters $\phi = \left\{\mathbf{Q}, \mathbf{P}, \sigma^2\right\}$ are obtained at $E\left(\frac{\partial L_c}{\partial \phi}\right) = 0$. To calculate the optimal parameters we solve (7), (8) and (9).

$$\mathbf{Q}\sum_{i,k} E\left(x_i x_i^T\right) = \sum_{i,k}\left(f_i^k - \mu\right)E(x_i) - \mathbf{P}E\left(y_k x_i^T\right) \tag{7}$$

$$\mathbf{P}\sum_{i,k} E\left(y_k y_k^T\right) = \sum_{i,k}\left(f_i^k - \mu\right)E(y_k) - \mathbf{Q}E\left(x_i y_k^T\right) \tag{8}$$

The optimal solution for the noise variance is represented as

$$\sigma^2 = \frac{1}{DN}\sum_{i,k}\left(f_i^k - \mu\right)^T\left(f_i^k - \mu - \mathbf{Q}E(x_i) - \mathbf{P}E(y_k)\right) \tag{9}$$

where D is the dimension of the mean feature and N denotes total number of training images. After learning parameters of the HFA model and corresponding low dimensional representation of feature space, we can use a regression technique to estimate the age of a query image.

3.2 Regression for Age Estimation

Let the local facial feature space $\mathcal{F}$ is represented as $\mathcal{F} = \left\{f_i : f_i \in \mathbb{R}^D\right\}_{i=1}^N$ where N is number of face images and D is dimension of the data. The ground truth age labels a_i are represented as $a = \{a_i : a_i \in \mathbb{N}\}_{i=1}^N$. Our objective is to learn a low dimensional discriminative manifold $\mathcal{G}$ embedded in $\mathcal{F}$ and subsequently a low dimensional aging feature $\left\{t_i : t_i \in \mathbb{R}^d\right\}_{i=1}^N$ with $d << D$. Once the aging subspace $\mathbf{P}$ is learned from the HFA training, our next task is to find low dimensional age discriminative features. The low dimensional aging features are obtained by projecting local feature onto a $D \times d$ projection matrix *i.e.* $T = \mathbf{P}^T F$ where $F = [f_1, f_2, \cdots, f_N] \in \mathbb{R}^{D \times N}$ and $\mathbf{P} = [p_1, p_2 \cdots, p_d]$. Having obtained the low-dimensional feature representations from the HFA model, we define the age estimation as a multiple linear regression problem in the low-dimensional manifold space. Various linear and nonlinear regression methods are available in the literature. To explore the discriminative power of the aging feature, we use multiple regression techniques such as support vector regression, Gaussian process regression [50] and hierarchical age estimation methods on the proposed aging feature.

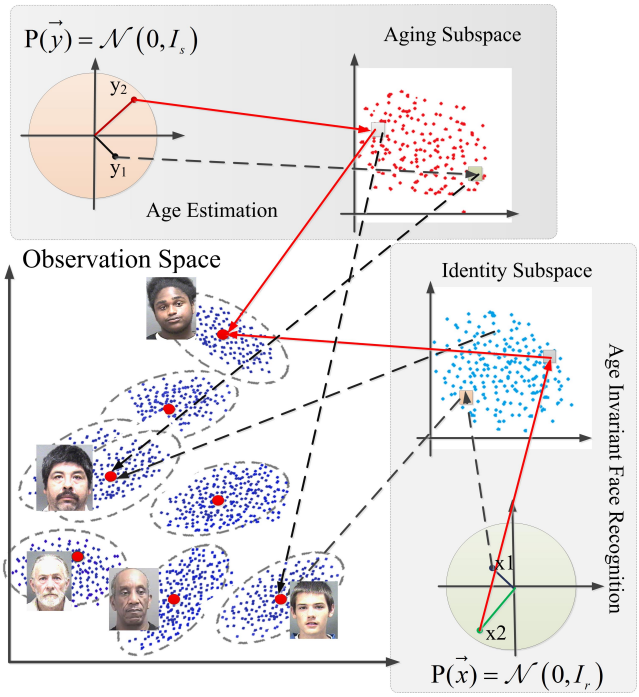

Fig. 1: Illustration of the proposed aging subspace learning scheme. Latent identity and age variables in (1) have a prior distribution $p(x) \sim \mathcal{N}(0, I_r)$ and $p(y) \sim \mathcal{N}(0, I_s)$. Circles in the observation space represent observed data for a person at different age. Identity factor is used for age invariant face recognition [12, 13, 24]. We propose aging subspace for age estimation.

3.3 Age Estimation Framework

Our age estimation framework mainly incorporates four modules: face preprocessing, feature extraction, discriminative aging subspace learning, and regression. In the training stage, face images undergo normalizations such as geometric alignments and illumination normalization (basically histogram equalization). Then, the histogram-of-oriented-gradient (HOG) [7] feature is extracted on each image. The extracted feature is of very high dimension. Next, the age manifold is learned from the hidden factor analysis model (presented in section 3.1), to map the features into a low-dimensional subspace. We further learn a regression function to fit the manifold data. In the test stage, an input face image goes through the same preprocessing and local feature extraction stage. Then the high dimensional local feature is transformed to the low-dimensional aging feature by projecting it on the learned aging manifold. Finally, the age of the input face image is estimated by fitting the regression on the learned low-dimensional aging feature.

4 Experiments

In this section, we first present evaluation metric and evaluation protocol used for performance analysis of the proposed work. Next, in the result section we first demonstrate the robustness of the proposed discriminative aging feature by using it along with the multiple regression schemes. Next, we perform age estimation experiments to to verify the effectiveness of the proposed approach and to compare the proposed work with the state-of-the-art methods.

Our approach of aging manifold based age estimation has been evaluated on MORPH-II [36] database. It contains 55,314 face images of 12,936 subjects in the age range 16 to 77 years. For training HFA model using MORPH-II dataset, we followed similar settings as in the recent works [12, 13, 24] where 12,000 images of 6,000 subjects are used for training HFA model. Remaining images from the MORPH-II dataset are used for age estimation experiments. Note that we first train HFA model to learn discriminative aging subspace and then extract aging feature for regression from the learned subspace. In the preprocessing, an input face image is first aligned and then cropped to size 200×160. The cropped face image is converted to a gray-scale image and then histogram equalized. We divide preprocessed

face image into overlapping patches and extract HOG feature from each patch. The extracted feature results into a very high dimension therefore we apply widely used principle component analysis (PCA) + linear discriminant analysis (LDA) prior to learn the HFA model. As per our settings the feature space dimension is 6000.

4.1 Evaluation Metrics and Protocol

Performance of age estimation techniques is assessed using two evaluation metrics, Mean Absolute Error (MAE) and Cumulative Score (CS). The smaller the MAE, the better the age estimation performance. MAE shows the average performance of the age estimation technique and is an appropriate measure when the training data has many missing images. MAE indicates the mean absolute error between the predicted result and the ground truth for testing set, and is given by,

$$MAE = \sum_{k=1}^{N} \frac{\left|\widehat{y_k} - y_k\right|}{N} \tag{10}$$

In addition to MAE, CS is another performance metric which computes the overall accuracy of the estimator and is defined as:

$$CS\,(k) = \frac{N_{e<k}}{N} \times 100\% \tag{11}$$

where $N_{e<k}$ is the number of test images for which the absolute error by the age estimation algorithm is not higher than k years. The higher the CS value, the better the age estimation performance. CS is a useful measure of performance in age estimation when the training dataset has samples at almost every age. However, in age estimation, due to imbalanced and skewed databases both MAE and CS are used for evaluation.

4.2 Result

4.2.1 Analysis of the Proposed Subspace

To demonstrate the effectiveness of the proposed manifold learning scheme, we perform regression on the aging features extracted from the discriminative aging subspace. We selected standard regression techniques such as support vector regression (SVR), orthogonal Gaussian process (OGP) regression [50] and hierarchical SVM-SVR [18,19] for age estimation. In SVR and SVM learning, LIBSVM [4] is used to evaluate the approaches. The SVM and SVR employ radial basis function kernel and their parameters were found by a grid search using 5-fold cross validation. Table 1 lists the

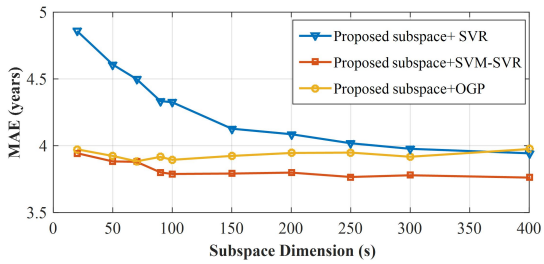

Fig. 2: MAE versus dimensionality of proposed aging subspace.

MAE and CS of three regression methods on the proposed aging manifold. We observed that hierarchical regression yields best performance on the proposed subspace. Note that the performance of proposed subspace along with SVR and OGP are also competent with the state-of-the-art age estimation methods.

Table 1: Performance of The Proposed Method with Different Regression Techniques

Regression Technique	**MAE**	**CS**
Proposed subspace+SVR	3.93	77
Proposed subspace+OGP	3.88	77
Proposed subspace+(SVM-SVR)	3.75	79

Furthermore, we have conducted experiments to demonstrate the effect of subspace dimensionality on MAE of three regression methods. Large parameter s implies more age information of a face image. We choose 400 as the upper limit for the dimension of learned mani-

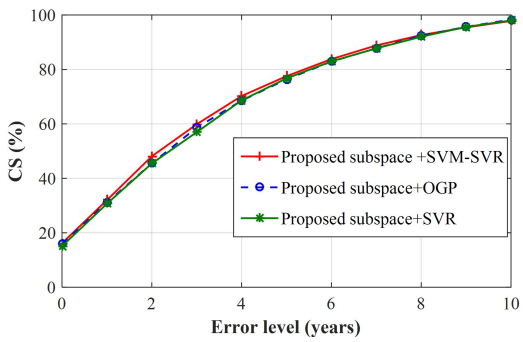

Fig. 3: Age estimation comparison of three regression methods on proposed aging manifold in terms of CS versus error levels.

Table 2: Effect of aging subspace and identity subspace dimension on computation time and accuracy

Aging Subspace Dimension (10)		**Aging Subspace Dimension (50)**		**Aging Subspace Dimension (100)**	
Identity Subspace Dimension	**MAE**	**Identity Subspace Dimension**	**MAE**	**Identity Subspace Dimension**	**MAE**
10	4.11	10	3.86	10	3.83
20	4.05	20	3.91	20	3.84
50	4.05	50	3.92	50	3.84
100	4.01	100	3.88	100	3.83
150	4.13	150	3.88	150	3.82
200	4	200	3.92	200	3.88
250	3.98	250	3.92	250	3.88
300	4.05	300	3.94	300	3.88

Aging Subspace Dimension (200)		**Aging Subspace Dimension (250)**		**Aging Subspace Dimension (300)**	
Identity Subspace Dimension	**MAE**	**Identity Subspace Dimension**	**MAE**	**Identity Subspace Dimension**	**MAE**
10	3.78	10	3.82	10	3.74
20	3.81	20	3.75	20	3.79
50	3.79	50	3.75	50	3.79
100	3.78	100	3.74	100	3.82
150	3.80	150	3.76	150	3.84
200	3.86	200	3.79	200	3.79
250	3.86	250	3.83	250	3.79
300	3.85	300	3.85	300	3.80

fold subspace. Note that the original feature space has dimension 6000. Fig. 2 compare MAE versus dimensionality of the proposed aging subspace on multiple regression methods. We observe in Fig. 2 that when subspace dimensionality reaches 300, the age discriminative information extracted by the aging manifold saturates. Comparative analysis in terms of CS is shown in Fig. 3.

To analyse effect of identity subspace dimension on the age estimation performance we conducted one more experiment, in which we fixed the aging subspace and varied identity subspace dimension and recorded MAE for that setting. The details of identity subspace dimension, aging subspace dimension and corresponding MAE are shown in Table 2. For the given aging subspace dimension, as far as age estimation is concerned, as already discussed the aging subspace dimension plays major role. However, the identity subspace dimension has hardly any role in the age estimation. It is clear from Table 2 that for any given aging subspace dimension, the MAE does not vary drastically with respect to the corresponding identity subspace dimension. Thus the MAE and identity subspace dimension are practically independent

4.2.2 Age Estimation Experiments

We also compare the proposed approach with the state-of-the-art age estimation algorithms such as appearance

Table 3: Comparison of The Proposed Age Estimation Method To Previous Works in Terms of MAE (Years) and CS (%)

Technique	**Facial Feature**	**MAE / CS**
AGES [10]	AAM	8.8 / 46%
OGP [50]	SIFT	3.92 / –
KPLS [15]	BIF	4.18 / –
Flexible Overlap [33]	$LPQ_{7\times7}$	5.86 / 68%
Guo et. al [17]	BIF	4.2 / –
LU and Tan [30]	Manifold of raw intensity	White: 5.2/- Black : 4.2/-
Han et. al. [18]	Demographic informative features	5.10 / –
CNN [20]	Deep Features	3.88 / –
CPNN [9]	Deep Features	4.87 / –
Proposed Method	Age discriminative features	**3.7 / 79%**

and age specific (AAS) [23], KPLS [15], hierarchical age estimation using SVM-SVR [19,27,33,41], and a few deep learning methods [21,44]. Table 3 shows the comparison study among proposed approach and various state-of-the-art approaches in terms of MAE and CS at 5 years error. Clearly, proposed method has a better MAE and CS than the competing approaches. This indicates that the proposed aging subspace provides discriminative aging feature and about 79% of the age predictions from it do not differ more than 5 years from the ground truth. Therefore, the proposed subspace learning from HFA model represents a robust aging manifold.

5 Conclusion

In this paper, we have proposed a probabilistic HFA approach to address the challenging problem of age estimation. The basic idea of the proposed model is to pursue a more robust aging manifold and corresponding aging feature descriptor from the age component of the HFA model. Extensive experiments conducted on MORPH Album II demonstrate that proposed subspace can extract discriminative aging features for age estimation.

Declarations

- **Ethical approval:** This study does not involve any human participants or animals performed by any of the authors.
- **Funding:** This research received no external funding.
- **Informed Consent Statement:** Not applicable.
- **Conflicts of Interest:** The authors declare no conflict of interest.
- **Author Contributions:** Conceptualization, methodology, investigation and coding, writing original draft preparation, M.S.; writing review and editing, M.S. and K.M.B.; supervision, K.M.B.

References

1. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: Application to face recognition. IEEE transactions on pattern analysis and machine intelligence **28**(12), 2037–2041 (2006)
2. Ahonen, T., Rahtu, E., Ojansivu, V., Heikkila, J.: Recognition of blurred faces using local phase quantization. In: 2008 19th international conference on pattern recognition, pp. 1–4. IEEE (2008)
3. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neural computation **15**(6), 1373–1396 (2003)
4. Chang, C.C., Lin, C.J.: Libsvm: a library for support vector machines. ACM transactions on intelligent systems and technology (TIST) **2**(3), 27 (2011)
5. Chang, K.Y., Chen, C.S., Hung, Y.P.: Ordinal hyperplanes ranker with cost sensitivities for age estimation. In: Computer vision and pattern recognition (cvpr), 2011 ieee conference on, pp. 585–592. IEEE (2011)
6. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Transactions on pattern analysis and machine intelligence **23**(6), 681–685 (2001)
7. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1, pp. 886–893. IEEE (2005)
8. Fu, Y., Huang, T.S.: Human age estimation with regression on discriminative aging manifold. IEEE Transactions on Multimedia **10**(4), 578–584 (2008)
9. Geng, X., Yin, C., Zhou, Z.H.: Facial age estimation by learning from label distributions. IEEE transactions on pattern analysis and machine intelligence **35**(10), 2401–2412 (2013)
10. Geng, X., Zhou, Z.H., Smith-Miles, K.: Automatic age estimation based on facial aging patterns. IEEE Transactions on pattern analysis and machine intelligence **29**(12), 2234–2240 (2007)
11. Geng, X., Zhou, Z.H., Zhang, Y., Li, G., Dai, H.: Learning from facial aging patterns for automatic age estimation. In: Proceedings of the 14th ACM international conference on Multimedia, pp. 307–316. ACM (2006)
12. Gong, D., Li, Z., Lin, D., Liu, J., Tang, X.: Hidden factor analysis for age invariant face recognition. In: Computer Vision (ICCV), 2013 IEEE International Conference on, pp. 2872–2879. IEEE (2013)
13. Gong, D., Li, Z., Tao, D., Liu, J., Li, X.: A maximum entropy feature descriptor for age invariant face recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 5289–5297 (2015)
14. Guo, G., Fu, Y., Dyer, C.R., Huang, T.S.: Image-based human age estimation by manifold learning and locally adjusted robust regression. IEEE Transactions on Image Processing **17**(7), 1178–1188 (2008)
15. Guo, G., Mu, G.: Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression. In: Computer vision and pattern recognition (cvpr), 2011 ieee conference on, pp. 657–664. IEEE (2011)
16. Guo, G., Mu, G.: Joint estimation of age, gender and ethnicity: Cca vs. pls. In: 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), pp. 1–6. IEEE (2013)
17. Guo, G., Mu, G., Fu, Y., Huang, T.S.: Human age estimation using bio-inspired features. In: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pp. 112–119. IEEE (2009)
18. Han, H., Otto, C., Jain, A.K.: Age estimation from face images: Human vs. machine performance. In: Biometrics (ICB), 2013 International Conference on, pp. 1–8. IEEE (2013)
19. Han, H., Otto, C., Liu, X., Jain, A.K.: Demographic estimation from face images: Human vs. machine performance. IEEE transactions on pattern analysis and machine intelligence **37**(6), 1148–1161 (2015)
20. Huerta, I., Fernández, C., Segura, C., Hernando, J., Prati, A.: A deep analysis on age estimation. Pattern Recognition Letters **68**, 239–249 (2015)
21. Huo, Z., Yang, X., Xing, C., Zhou, Y., Hou, P., Lv, J., Geng, X.: Deep age distribution learning for apparent age estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 17–24 (2016)

22. Iqbal, M.T.B., Shoyaib, M., Ryu, B., Abdullah-Al-Wadud, M., Chae, O.: Directional age-primitive pattern (dapp) for human age group recognition and age estimation. IEEE Transactions on Information Forensics and Security **12**(11), 2505–2517 (2017)
23. Lanitis, A., Taylor, C.J., Cootes, T.F.: Toward automatic simulation of aging effects on face images. IEEE Transactions on Pattern Analysis and Machine Intelligence **24**(4), 442–455 (2002)
24. Li, H., Zou, H., Hu, H.: Modified hidden factor analysis for cross-age face recognition. IEEE Signal Processing Letters **24**(4), 465–469 (2017)
25. Li, Z., Gong, D., Li, X., Tao, D.: Aging face recognition: A hierarchical learning model based on local patterns selection. IEEE Transactions on Image Processing **25**(5), 2146–2154 (2016)
26. Li, Z., Park, U., Jain, A.K.: A discriminative model for age invariant face recognition. IEEE transactions on information forensics and security **6**(3), 1028–1037 (2011)
27. Liang, Y., Wang, X., Zhang, L., Wang, Z.: A hierarchical framework for facial age estimation. Mathematical Problems in Engineering **2014** (2014)
28. Ling, H., Soatto, S., Ramanathan, N., Jacobs, D.W.: A study of face recognition as people age. In: Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, pp. 1–8. IEEE (2007)
29. Lowe, D.G.: Object recognition from local scale-invariant features. In: Proceedings of the seventh IEEE international conference on computer vision, vol. 2, pp. 1150–1157. IEEE (1999)
30. Lu, J., Tan, Y.P.: Ordinary preserving manifold analysis for human age and head pose estimation. IEEE Transactions on Human-Machine Systems **43**(2), 249–258 (2013)
31. Moon, T.K.: The expectation-maximization algorithm. IEEE Signal processing magazine **13**(6), 47–60 (1996)
32. Park, U., Tong, Y., Jain, A.K.: Age-invariant face recognition. IEEE transactions on pattern analysis and machine intelligence **32**(5), 947–954 (2010)
33. Pontes, J.K., Britto Jr, A.S., Fookes, C., Koerich, A.L.: A flexible hierarchical approach for facial age estimation based on multiple features. Pattern Recognition **54**, 34–51 (2016)
34. Prince, S.J., Elder, J.H.: Probabilistic linear discriminant analysis for inferences about identity. In: Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, pp. 1–8. IEEE (2007)
35. Prince, S.J., Elder, J.H., Warrell, J., Felisberti, F.M.: Tied factor analysis for face recognition across large pose differences. IEEE Transactions on pattern analysis and machine intelligence **30**(6), 970–984 (2008)
36. Ricanek, K., Tesafaye, T.: Morph: A longitudinal image database of normal adult age-progression. In: Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on, pp. 341–345. IEEE (2006)
37. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. science **290**(5500), 2323–2326 (2000)
38. Sawant, M., Addepalli, S., Bhurchandi, K.: Age estimation using local direction and moment pattern (ldmp) features. Multimedia Tools and Applications **78**(21), 30419–30441 (2019)
39. Sawant, M., Bhurchandi, K.: Hierarchical facial age estimation using gaussian process regression. IEEE Access **7**, 9142–9152 (2019)
40. Tenenbaum, J.B., De Silva, V., Langford, J.C.: A global geometric framework for nonlinear dimensionality reduction. science **290**(5500), 2319–2323 (2000)
41. Thukral, P., Mitra, K., Chellappa, R.: A hierarchical approach for human age estimation. In: Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, pp. 1529–1532. IEEE (2012)
42. Wang, J., Shang, Y., Su, G., Lin, X.: Age simulation for face recognition. In: Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, vol. 3, pp. 913–916. IEEE (2006)
43. Wang, S., Tao, D., Yang, J.: Relative attribute svm+ learning for age estimation. IEEE transactions on cybernetics **46**(3), 827–839 (2016)
44. Wang, X., Guo, R., Kambhamettu, C.: Deeply-learned feature for age estimation. In: Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on, pp. 534–541. IEEE (2015)
45. Wen, Y., Li, Z., Qiao, Y.: Latent factor guided convolutional neural networks for age-invariant face recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4893–4901 (2016)
46. Weng, R., Lu, J., Yang, G., Tan, Y.P.: Multi-feature ordinal ranking for facial age estimation. In: 2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), pp. 1–6. IEEE (2013)
47. Xiao, B., Yang, X., Xu, Y., Zha, H.: Learning distance metric for regression by semidefinite programming with application to human age estimation. In: Proceedings of the 17th ACM international conference on Multimedia, pp. 451–460 (2009)
48. Xu, C., Liu, Q., Ye, M.: Age invariant face recognition and retrieval by coupled auto-encoder networks. Neurocomputing **222**, 62–71 (2017)
49. Zhang, Y., Yeung, D.Y.: Multi-task warped gaussian process for personalized age estimation. In: Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pp. 2622–2629. IEEE (2010)
50. Zhu, K., Gong, D., Li, Z., Tang, X.: Orthogonal gaussian process for automatic age estimation. In: Proceedings of the 22nd ACM international conference on Multimedia, pp. 857–860. ACM (2014)