A Study on Human Age Estimation Under Facial Expression Changes

Guodong Guo and Xiaolong Wang
West Virginia University
Lane Dept. of CSEE, Morgantown, WV
guodong.guo@mail.wvu.edu, xwang5@mix.wvu.edu

Abstract

In this paper, we study human age estimation in face images under significant expression changes. We will address two issues: (1) Is age estimation affected by facial expression changes and how significant is the influence? (2) How to develop a robust method to perform age estimation undergoing various facial expression changes? This systematic study will not only discover the relation between age estimation and expression changes, but also contribute a robust solution to solve the problem of cross-expression age estimation. This study is an important step towards developing a practical and robust age estimation system that allows users to present their faces naturally (with various expressions) rather than constrained to the neutral expression only. Two databases originally captured in the Psychology community are introduced to Computer Vision, to quantitatively demonstrate the influence of expression changes on age estimation, and evaluate the proposed framework and corresponding methods for cross-expression age estimation.

1. Introduction

Human age estimation (HAE) has recently become an active research topic in Computer Vision [3, 11], because of many potential applications in practice, such as business intelligence, security, and human-centered image understanding. However, age estimation is still a very challenging problem, and lots of issues have not been well addressed. For instance, it is not clear yet about the relation between age estimation and facial expression changes. Human subjects can perform various expressions naturally and very often in daily lives. To develop a practical and robust age estimation system, we need to study if age estimation is influenced by expression changes, how significant is the influence, and if a solution could be developed to solve the problem caused by facial expressions. Existing works on age estimation, e.g., [19, 5, 10, 7, 24, 4], are mainly based on neutral faces. The popular aging databases, e.g., FG-NET [2] and MORPH [12] contain mainly frontal-view, neutral faces, although there are some various in illumination, pose, and expression. To perform a systematic study on age estimation under expression changes, we need to use databases with clear ground truth labels of both age and expression.

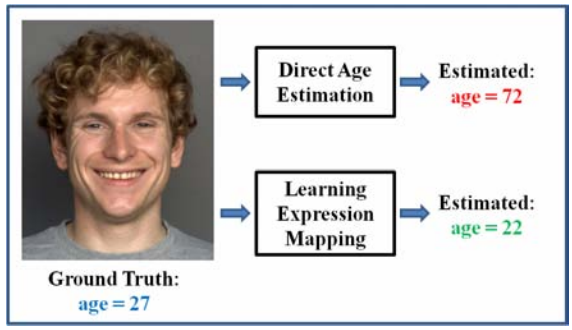

Figure 1. An example face to show the influence of expression on age estimation. A young, happy face (age = 27 years) can have an estimated age of 72 based on one of the state-of-the-art methods. The happy expression causing wrinkles in face regions can "fool" the age estimator. A well-designed method in this paper can get an age estimation of 22 years, close to the ground truth.

In this paper, we will systematically study the problem of age estimation under facial expression changes. The critical need to execute this study is shown in Figure 1. We will also propose a robust solution to address the problem of age estimation across expression changes. Our major contributions in this paper are described as follows.

1.1. Major Contributions

1. We show that age estimation is influenced by expression changes, based on a *quantitative evaluation* on two databases newly introduced to Computer Vision;

2. We propose a new framework for age estimation across expression changes. The concept of *learning expression correlations* is completely new to age estimation;

3. We explore appropriate methods to implement the framework and demonstrate a good performance towards *cross-expression age estimation*.

978-1-4673-1228-8/12/$31.00 ©2012 IEEE

The paper is organized as follows. In Section 2, we present a new framework for age estimation across expression changes. Then we describe methods for learning expression correlations, discriminative mapping, and the aging function in Sections 3, 4, and 5, respectively. In Section 6, we introduce two databases that are used for our research. Finally, we perform various experimental evaluations and draw conclusions.

2. A New Framework

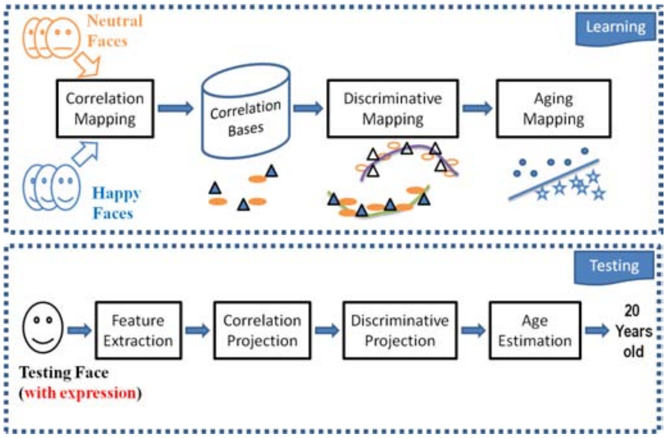

Figure 2. The proposed *framework* for age estimation under expression changes in face images. To be concise, only the happy expression is used, but the framework can be applied to other expressions too. Our experimental validations have used all available expressions.

To deal with the influence of expression changes (see quantitative evaluations of the influence in Section 7), and make it possible to perform age estimation across expression changes, we propose a new framework shown in Figure 2. First, we propose to learn the *correlations* between neutral and other expressions, e.g., happy or sad. The correlations are learned from pairs of expressions performed by same individuals, otherwise, the correlations may not exist. The faces will be projected to the learned correlation bases to form new features. Second, we propose to perform a discriminative mapping that can make the patterns of the same age closer, while different ages separated further. Our argument is that the correlation mapping can "move" aging patterns of different expressions closer, but not necessarily making patterns at the same age closer. Then, we learn the aging mapping or aging function as traditional approaches [4]. We prefer to use classifiers in this work. Please note that in correlation mapping, one may use face images directly or extracted features from the face images. We found that using features are much better than raw images.

In testing, features are extracted first for a given face image. Then the features are projected onto the correlation bases and discriminative bases. Finally age estimation is executed using the aging function. In the following, we will explore appropriate methods to implement the new framework and perform experimental validations.

3. Learning Expression Correlations

We propose to learn the *correlations* between neutral and other expressions, in order to deal with the influence of expression changes. The goal is to use the learned correlation bases to map the aging features extracted from faces with other expressions to new features that are "*virtually*" extracted from neutral faces at the same age. The challenge is that the correlation mapping of expressions needs to preserve the aging information. The idea of learning expression correlations has two advantages: (1) There is no need to synthesize neutral faces from other expressions. Synthesis of facial expressions itself is a hard problem [8]. Misalignment of images and unnecessary facial smoothing may lose aging details in the synthesized faces. (2) The computation is faster than synthesizing new expression images.

Different methods could potentially be used to learn the expression correlations. In our approach, we apply the PLS algorithm to learn the correlation, which places the PLS method into a new problem to exploit its new capability.

3.1. Partial Least Squares

Partial Least Squares (PLS) methods model relationships between sets of observed variables by means of latent variables [22, 14, 13]. The underlying assumptions of all PLS methods is that the observed data is generated by a system which is driven by a small number of latent variables. The projection of the observed data onto the latent variables has been shown to be a powerful technique when observed variables are highly correlated.

Suppose there are two data sets or blocks of variables that are correlated, denoted by $\mathcal{X} \subset \mathcal{R}^N$ and $\mathcal{Y} \subset \mathcal{R}^M$, respectively, where N and M are the dimensions of the two spaces. PLS models the relations between these two blocks by means of score vectors. Given n data samples, PLS decomposes the $(n \times N)$ matrix of zero-mean variables $\mathbf{X}$ and the $(n \times M)$ matrix of zero-mean variables $\mathbf{Y}$ into the form

$$\begin{aligned} \mathbf{X} &= \mathbf{T}\mathbf{P}^T + \mathbf{E} \\ \mathbf{Y} &= \mathbf{U}\mathbf{Q}^T + \mathbf{F} \end{aligned} \quad (1)$$

where the $\mathbf{T}$ and $\mathbf{U}$ are $(n \times p)$ matrices of the p extracted score vectors (components, latent vectors), the $(N \times p)$ matrix $\mathbf{P}$ and the $(M \times p)$ matrix $\mathbf{Q}$ represent matrices of loadings, and the $(n \times N)$ matrix $\mathbf{E}$ and the $(n \times M)$ matrix $\mathbf{F}$ are the matrices of residuals. The classical PLS method is based on the nonlinear iterative partial least squares (NIPALS) algorithm [21], finds weight vectors $\mathbf{w}, \mathbf{c}$ such that

$$\begin{aligned} [cov(\mathbf{t}, \mathbf{u})]^2 &= [cov(\mathbf{Xw}, \mathbf{Yc})]^2 \\ &= max_{|r|=|s|=1} [cov(\mathbf{Xr}, \mathbf{Ys})]^2 \end{aligned} \quad (2)$$

where $cov(\mathbf{t}, \mathbf{u}) = \frac{\mathbf{t}^T \mathbf{u}}{n}$ denotes the sample covariance between the score vectors $\mathbf{t}$ and $\mathbf{u}$. The NIPALS algorithm starts with random initialization of the $\mathcal{Y}$-space score vector $\mathbf{u}$ and repeats a sequence of iterations until convergence [21].

The linear PLS models can have variants based on the deflation difference [14]. Most PLS models assume that (1) the score vectors $\mathbf{t}_i$, $i = 1, \cdots, p$, are good predictors of $\mathbf{Y}$, and (2) a linear *inner relation* between the score vectors $\mathbf{t}$ and $\mathbf{u}$ exists; that is

$$\mathbf{U} = \mathbf{T}\mathbf{D} + \mathbf{H} \tag{3}$$

where $\mathbf{D}$ is a $(p \times p)$ diagonal matrix and $\mathbf{H}$ is the matrix of residuals. Combining Eqns. (1) and (3), we obtain

$$\mathbf{Y} = \mathbf{T}\mathbf{D}\mathbf{Q}^T + (\mathbf{H}\mathbf{Q}^T + \mathbf{F}) \tag{4}$$

and this defines the *linear PLS regression* model

$$\mathbf{Y} = \mathbf{T}\mathbf{C}^T + \mathbf{F}^* \tag{5}$$

where $\mathbf{C}^T = \mathbf{D}\mathbf{Q}^T$ denotes the $(p \times M)$ matrix of regression coefficients and $\mathbf{F}^* = \mathbf{H}\mathbf{Q}^T + \mathbf{F}$ is the residual matrix.

The PLS models were originally derived for regression problems [21, 22, 14], but can be adapted to classification using a similar form as Eqn. (5). The difference is the $\mathbf{Y}$ matrix that can be changed to encode the class membership. Recently, the PLS method has been adapted to solve computer vision problems successfully, e.g., pedestrian detection [16], face recognition [15, 18], and age estimation with neutral expression only [5]. Most of the previous approaches focus on the exploration of feature dimensionality reduction by the PLS [16, 15, 5], using image labels as the second block of variables $\mathbf{Y}$. Sharma and Jacobs [18] proposed to use raw images for both sets of variables in PLS to deal with multi-modality face recognition, e.g., photo and sketch, but their work has nothing to do with facial expression or age estimation. Here we want to adapt the PLS method to learn expression correlations for the purpose of *cross-expression age estimation*, which is a novel exploitation of the PLS. More importantly, we will show that the PLS itself cannot solve our problem, although it is a critical component in the whole framework that we proposed.

4. Discriminative Mapping

The correlation mapping can build the "relations" between different expressions, but it does not necessarily consider the pattern distributions over different ages. For age estimation, we hope the mapped patterns to have "close" distributions if they belong to the same age, but become "separated" further if they belong to different ages. The PLS-based expression correlation cannot have this capability, thus we propose to have a discriminative learning step after correlations. The idea of *correlation-discrimination pairing* has not been investigated in any previous approach to the best of our knowledge. It might be explored broadly in other vision problems.

In our discriminative mapping, we use the Marginal Fisher Analysis (MFA) [23], which is a supervised manifold learning algorithm with Fisher criterion. It constructs the within-class graph $\mathcal{G}_{\mathrm{w}}$ and between-class graph $\mathcal{G}_{\mathrm{b}}$ considering both discriminant and geometrical structure in the data. Define the within-class affinity weight $s_{ij}^{(\mathrm{w})} = 1$ when $\mathbf{x}_i$ and $\mathbf{x}_j$ are k nearest neighbors of each other with the same class label, otherwise $s_{ij}^{(\mathrm{w})} = 0$. Define symmetric matrix $\mathbf{S}_{\mathrm{w}}(i,j) = s_{ij}^{(\mathrm{w})}$, diagonal matrix $\mathbf{D}_{\mathrm{w}}(i,i) = \sum_j s_{ji}^{(\mathrm{w})}$, and Laplacian matrix $\mathbf{L}_{\mathrm{w}} = \mathbf{D}_{\mathrm{w}} - \mathbf{S}_{\mathrm{w}}$. Similarly, define the between-class affinity weight $s_{ij}^{(\mathrm{b})} = 1$ when $\mathbf{x}_i$ and $\mathbf{x}_j$ are k nearest neighbors of each other with different class labels, otherwise $s_{ij}^{(\mathrm{b})} = 0$. Thus $\mathbf{S}_{\mathrm{b}}$, $\mathbf{D}_{\mathrm{b}}$ and $\mathbf{L}_{\mathrm{b}}$ are obtained. The objective of MFA is

$$\mathbf{p} = \underset{\mathbf{p}}{\mathrm{argmin}} \frac{\mathbf{p}^T \mathbf{X} \mathbf{L}_{\mathrm{w}} \mathbf{X}^T \mathbf{p}}{\mathbf{p}^T \mathbf{X} \mathbf{L}_{\mathrm{b}} \mathbf{X}^T \mathbf{p}}$$

Here we use the MFA for discriminative mapping of the aging patterns after the correlation mapping of two different expressions. It is a novel endeavor to explore the MFA method for cross-expression age estimation. Further, the order is important to do discriminative mapping. It should be performed after correlation mapping. Switching the order between them will result in performance drop (we verified this experimentally, but it is not shown here).

5. Aging Function Learning and Features

After discriminative mapping of the aging patterns, we will learn the aging function that maps the transformed aging patterns to age labels, e.g., 25 or 30. The aging function can be learned using classifiers, regressors, or a combination of both [4]. Here we use the Support Vector Machine (SVM) classifier [20] to learn the aging function.

For facial aging image representation, we adopted one of the state-of-the-art methods, called Biologically-Inspired Features (BIF) [7]. It is based on a feed-forward structure of signal encoding in vision cortex [17], containing both simple and complex units, denoted by $S1$ and $C1$. Raw face images cannot be used for aging representation based on our evaluation (not shown here). The BIF representation is useful and invariant to some small variations, but it is not invariant to significant expression changes in our new problem, as will be demonstrated in the experiments.

6. Two Databases

The FG-NET [2] and MORPH [12] aging databases are not proper for our study. These two databases contain faces

Figure 3. Age estimation in face images with different expressions. Each individual's faces have significant expression changes (see each row) but with the same ground truth age (46 and 72 years old, respectively). The question is: Can we develop a method to estimate age no matter how the expression changes?

Table 1. Two aging databases (I: Lifespan and II: FACES) with facial expressions, showing the number of faces in four age groups.

D.	Exp.	Age Group				Total faces
		18-29	30-49	50-69	70-94	
I	Neutral	223	76	133	158	844
	Happy	142	19	48	45	
II	6 Exps.	51	35	31	54	1026

mainly with the neutral expression, although there are a small number of faces with some variations in illumination, expression and pose. Both databases have no labeling about facial expressions.

In computer vision, there is few aging database with various facial expressions (and expression labels). From the psychology society, we found two databases that contain both aging and facial expressions with ground truth labels, called the Lifespan [9] and FACES [1] databases. We introduce these two databases to the computer vision community, and use both of them for our study.

The Lifetime database [9] contains 844 frontal face images with neutral and happy expressions. The age range is from 18 to 94 years. The database captured 590 individuals with neutral faces, and partial of them have the happy expression. The FACES database [1] contains 171 individuals with 6 expressions (neutral, sad, disgust, fear, angry, and happy) for each person in frontal view. Some example faces are shown in Fig. 3. The total number of face images is 1026 (actually 2052 in two sets, but the two sets are almost the same. Thus we only used one set). The Lifespan [9] contains more individuals than FACES [1], but less number of expressions. We will use both databases for our study. See Table 1 for more details about the two databases.

7. Experiments

We perform various age estimation experiments on the two aging and expression databases just introduced. We want to show the age estimation results in different cases *quantitatively*, measured by the mean absolute error (MAE). The MAE is defined by $\sum_{k=1}^{N} |\hat{g}_k - g_k|/N$, where g_k is the ground truth age for the test image k, $\hat{g}_k$ is the estimated age, and N is the number of test images.

We use five-fold cross validation in our experiments. When age estimation is performed using the same facial expression, it is a standard five-fold cross validation defined in machine learning. When age estimation is performed across expressions, e.g., from neutral to happy, which means the ages of neutral faces are used for training, while the ages of happy faces are used for testing, the five-fold cross validation is then slightly different. As illustrated in Figure 4, we want to avoid the identity overlap in our experiments. To achieve this, we "align" the same identity to appear in the same partition of the five folds. For example, if person A's neutral face appears in N1, his/her happy face will appear in H1 correspondingly. In cross-expression age estimation, if sets N1 to N4 are used for training, H5 will be used for testing. As a result, the identities in H5 will never appear in the training set (neither N1 to N4 nor H1 to H4).

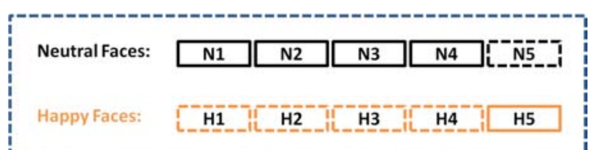

Figure 4. Five-fold cross validation scheme for age estimation with the same or different expressions.

In all our experiments, there is no overlap of individuals in any training and test sets. As a general rule in machine learning, it is important to design the experiment with no data overlapping in training and testing. Further, we avoid to use the same identity in both the training and test sets.

7.1. Age Estimation under the Same Expression

First, let us look at the age estimation results using the same expression. In this case, the training and test sets have faces with the same expression but different identities. The BIF [7] is one of the state-of-the-art representation for aging images. It can be further combined with manifold learning techniques to improve age estimation [6]. Here we adopt the BIF representation with the MFA method [23] for our study on a new problem using new databases. By performing a standard five-fold cross validation test on two databases, i.e., the FACES and Lifespan, we got the MAEs for age estimation shown in Table 2. There are six expressions in FACES while two expressions in Lifespan. From Table 2, one can see that the BIF+MFA has lower errors than the BIF only. More importantly, the MAEs for Neutral→Neutral, i.e., using both neutral faces for training and test, are smaller than using other expressions, e.g., Happy→Happy. These kinds of results are observed in both databases. To understand how big the differences are, we compute the MAE increase rates for all other expressions compared with the Neutral→Neutral case, using the BIF+MFA representation (last column in Table 2). The error increase rates range from 26.8% to 50.4% in FACES and 21.7% in Lifespan. This result tells that the age estimation errors will be much larger using various facial expressions rather than the neutral. The neutral expression gives significantly lower errors in both databases. This empirical evidence uncover the fact that facial expressions have a significant impact on age estimation. Our interpretation is that aging and emotional expression are "subtle" features in faces, and expression changes can cause aging details to change, e.g., wrinkles. One specific example is shown in Figure 1. Another interpretation is that the effects of expression changes on aging details vary with people. In other words, the expression changes are not "stable" for age estimation. It also indicates that the neutral expression is more stable than other expressions for age estimation.

7.2. Age Estimation under Expression Changes

Second, we would like to empirically study the problem of cross-expression age estimation. This is to further our study on the influence of facial expression changes on age estimation, and also verify if we can develop a method to solve the problem of age estimation across expression changes.

The experimental results on two databases are shown in Table 3. We consider the crossing situations from neutral to all other expressions that are available in the databases. The practical significance is to verify if one can use the neutral expression to train an age estimator, and then apply to any other facial expressions. To make it complete, we also consider the reverse situations, i.e., from other expressions to neutral.

Table 2. Age estimation using the same expression: Mean Abstract Error or MAE (in yrs.). The percentages in parentheses measure the amount of error increases in age estimation when different expressions are used compared to the neutral.

Database	Training	Test	Representation	
			BIF	BIF+MFA
FACES	Neutral	Neutral	9.50	8.14
	Happy	Happy	10.69	10.32(26.8%)
	Disgust	Disgust	13.23	12.24(50.4%)
	Fearful	Fearful	12.65	10.73(31.8%)
	Sad	Sad	10.78	10.66(31.0%)
	Angry	Angry	13.26	10.96(34.6%)
Lifespan	Neutral	Neutral	8.93	6.05
	Happy	Happy	10.75	7.36(21.7%)

In each expression crossing, we keep the same training examples, but change the testing faces corresponding to the target expression. A five-fold cross validation scheme was taken. In Table 3, columns 1 and 2, denoted by C_1 and C_2, show the results based on the BIF and BIF+MFA for aging face representation, and the SVM for aging function learning. In these two columns, the results are based on a straight forward age estimation, without learning the correlation between different expressions, e.g., neutral and happy. One can see that the MAEs are pretty large, if compared to the MAEs for age estimation under the same expressions or using the neutral only (see Table 2). This cross-expression evaluation further indicates that age estimation is influenced significantly by facial expression changes.

Next, we want to evaluate our proposed framework for cross-expression age estimation. Our expectation is that cross-expression age estimation could be dealt with through designing an appropriate framework. The key idea is to "correlate" the other expressions with the neutral, when performed by the same person at the same age. The correlation model can be learned from an assemble of facial expression pairs.

To validate this idea, we will learn a "correlation mapping" for each pair of expressions between neutral and the others. Then the other expressions can be "mapped" to neutral before age estimation. In this way, we can do age estimation across expression changes. To learn the correlation mapping, we use the PLS method.

We apply the PLS method to the BIF face representation. The results are shown in column 3 of Table 3. One can see that the MAEs are reduced significantly in many cases. For example, it is reduced from 16.57 down to 12.11 years for Neutral→Happy on FACES. The MAE reduction rate is 26.9%. However, the reduced MAE is still larger than the case of Neutral→Neutral. So the observation is that the PLS based correlation mapping can lower the age estimation errors than without correlation learning, but the error reduction is still not enough.

Through analyzing the age estimation problem, we believe that learning the correlation between expressions can help cross-expression age estimation, but the correlation itself is not sufficient to solve the overall problem of age estimation across expression changes. The underlying reason is that the correlation mapping can make different expressions "closer" in terms of the data distribution, but it does not consider to make the distributions of different ages as "separated" as possible. That is, to make the projected new features as much as "discriminative" for age estimation. Our goal is to make the performance of cross-expression age estimation as close as possible to the age estimation using the neutral expression.

To make the patterns of different ages separated far away after the PLS correlations, while making patterns of the same age close, we apply a supervised manifold learning technique to achieve this goal. Specifically, we choose the MFA method for our problem, although other manifold learning methods could also be used. The main purpose is to demonstrate the idea and the proposed framework. The results are shown in columns 4 and 5 of Table 3. Both columns are based on the BIF+PLS+MFA representations, but with some difference in details.

In applying the MFA (and later the SVM), we have different schemes to use the training data. Let us use an example to interpret this. Suppose we are dealing with age estimation across expression change: Neutral→Happy. In supervised manifold learning, we may use the age labels of faces for both expressions or just the neutral expression. In SVM training, we can also have two choices, using both expressions or just use the neutral after applying MFA. We want to see the difference between the two schemes experimentally. Let "*-1" denote the use of neutral expression only in MFA (and SVM) learning; while "*-2" denote the use of both expressions in learning. Remember that in correlation learning, both expressions are used. The results are shown in column 4 and 5 of Table 3. From the table, one can observe that the MAEs are reduced significantly using either scheme. The second scheme is better than the first in more cases. The lowest MAE among all approaches is shown in bold font for each case (see each row in the table).

To measure the significance quantitatively, we compute the *MAE reduction rate* and display in the last column of Table 3. We select the minimum error from columns 4 and 5 and compare with the minimum error from columns 1 and 2. The error reduction rates are quite significant, ranging from 16.8% to 47.7%, depending on which expression is used. Further, the reduced MAEs are comparable to the neutral expression based age estimation in several cases. For instance, the MAE is reduced to 8.66 years for Neutral→Happy on FACES, close to the 8.14 for Neutral→Neutral (see Table 2). This result demonstrates that as the first attempt, we can develop a framework to perform cross-expression age estimation. We also notice that there are some hard cases where MAEs are still higher than Neutral→Neutral, including Neutral→Disgust and Neutral→Sad. This indicates that further efforts are needed in those cases. Some parameters: We set 50 latent variables for PLS, about 30 dimensions for MFA bases, and RBF kernel for SVM.

Age estimation across expressions is of great value in practice. First, in training age estimators, we may not need to collect all possible expressions for each individual. Second, in testing, the users can be flexible and natural to present their faces, rather than being constrained to perform the neutral expression only.

7.3. Effect of Identity on Age Estimation

In all our age estimation experiments, we deliberately *avoid* the identity overlap in the training and test sets. It means that if a person's face appears in training, his/her face cannot be in the test set, no matter what the expressions are. Now we show why we emphasize this point in the paper.

In this experiment, we just use the BIF representation and SVM (for aging function learning), without using any other methods. We use the same individuals but with different expressions for training and testing, respectively, on the Lifespan database. Individuals with neutral expression is used for training the age estimator, while their faces with happy expression is used for testing. The result is shown in Table 4. The obtained MAE is 4.90 years, which is even smaller than our age estimation evaluation for Neutral→Neutral on the Lifespan: 8.93 years using BIF, and 6.05 years using the BIF+MFA (see Table 2). We also switched the training and test sets, and obtained an MAE of 4.06 years, which is also very small.

Table 4. MAEs (in yrs.) of small values, showing the effect of identity on age estimation using the Lifespan database.

Rep.	Case	Neu.→Happy	Happy→Neu.
BIF	w/. identity	4.90	4.06

From this experiment, we show that the identity can have a large impact on age estimation. Even if the expressions are changed from neutral to happy or vice versa, the identity information is kept, which can "help" to deliver a "good" age estimation result with small errors. Actually, this is more like an identity recognition rather than age estimation. Indirectly, this demonstrates that facial expression changes may

Table 3. Age estimation under facial expression changes: MAE (in yrs.). The percentages in last column are the MAE reduction rates using the new framework (Col. 4 or 5) compared to a direct estimation (Col. 1 or 2). "B+P+M" indicates "BIF+PLS+MFA." See text for "-1,-2."

Database	Training	Test	Representation					MAE Reduction Rate
			BIF	BIF+MFA	BIF+PLS	B+P+M-1	B+P+M-2	$\frac{\min(C_1,C_2)-\min(C_4,C_5)}{\min(C_1,C_2)}$
FACES	Neutral	Happy	16.57	16.73	12.11	9.18	**8.66**	47.7%
		Disgust	16.34	15.78	14.09	**10.86**	11.87	31.2%
		Fearful	14.71	14.21	10.80	9.64	**8.57**	39.7%
		Sad	14.47	13.59	12.63	10.55	**10.54**	22.4%
		Angry	14.17	13.81	14.08	11.68	**9.75**	29.4%
	Happy	Neutral	15.24	11.23	9.51	9.21	**8.11**	27.8%
	Disgust		12.55	11.92	10.00	10.39	**8.57**	28.1%
	Fearful		14.39	11.12	9.92	**9.25**	9.28	16.8%
	Sad		12.16	10.95	9.98	9.24	**8.66**	20.9%
	Angry		14.21	12.15	9.36	**8.26**	8.66	32.0%
Lifespan	Neutral	Happy	15.71	11.48	10.63	7.83	**6.34**	44.8%
	Happy	Neutral	11.42	8.81	7.47	7.08	**6.19**	29.7%

have less influence on identity recognition but more significant influence on age estimation, when compared with other results in the paper. This simple experiment tells that we should pay attention to avoiding identity overlap in age estimation.

8. Conclusions

We have studied the problem of age estimation under facial expression changes. A quantitative evaluation has shown the significant influence of expressions on age estimation. A new framework has been proposed to perform age estimation across expression changes. We have shown that cross-expression age estimation can obtain a reasonably good performance when learning expression correlations and discriminative mapping are executed before the aging function learning. Two databases with age and expression from the Psychology society have been introduced to Computer Vision and used to validate our studies. The promising result may inspire further research.

References

[1] N. Ebner, M. Riediger, and U. Lindenberger. Faces - A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, 42(1):351–362, 2010.

[2] FGNET. The fg-net aging database. In *http://www.fgnet.rsunit.com/*, 2002.

[3] Y. Fu, G.-D. Guo, and T. S. Huang. Age synthesis and estimation via faces: A survey. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32(11):1955–1976, 2010.

[4] G.-D. 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*, 17(7):1178–1188, 2008.

[5] G.-D. Guo and G. Mu. Simultaneous dimensionality reduction and human age estimation via kernel partial least squares regression. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 657–664, 2011.

[6] G.-D. Guo, G. Mu, Y. Fu, C. Dyer, and T. S. Huang. A study on automatic age estimation on a large database. In *IEEE International Conference on Computer Vision*, pages 1986–1991, 2009.

[7] G.-D. Guo, G. Mu, Y. Fu, and T. S. Huang. Human age estimation using bio-inspired features. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 112–119, 2009.

[8] D. Huang and F. D. la Torre. Bilinear kernel reduced rank regression for facial expression synthesis. In *ECCV*, 2010.

[9] M. Minear and D. Park. A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers*, 36(4):630–633, 2004.

[10] B. Ni, Z. Song, and S. Yan. Web image mining towards universal age estimator. In *ACM Multimedia*, 2009.

[11] N. Ramanathan, R. Chellappa, and S. Biswas. Age progression in human faces : A survey. *Visual Languages and Computing*, 2009.

[12] K. Ricanek and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. In *IEEE Conf. on AFGR*, pages 341–345, 2006.

[13] R. Rosipal. Nonlinear partial least squares: An overview. In H. Lodhi and Y. Yamanishi, editors, *Chemoinformatics and Advanced Machine Learning Perspectives: Complex Computational Methods and Collaborative Techniques*, pages 169–189. ACCM, IGI Global, 2011.

[14] R. Rosipal and N. Kramer. Overview and recent advances in partial least squares. In C. Saunders, M. Grobelnik, S. Gunn, and J. Shawe-Taylor, editors, *Subspace, Latent Structure and Feature Selection Techniques*, pages 34–51. Springer, 2006.

[15] W. Schwartz, H. Guo, and L. Davis. A robust and scalable approach to face identification. In *ECCV*, 2010.

[16] W. Schwartz, A. Kembhavi, D. Harwood, and L. Davis. Human detection using partial least squares analysis. In *ICCV*, 2009.

[17] T. Serre, L. Wolf, and T. Poggio. Object recognition with features inspired by visual cortex. In *IEEE Conf. on Comput. Vision and Pattern Recognit.*, 2005.

[18] A. Sharma and D. Jacobs. Bypassing synthesis: PLS for face recognition with pose, low-resolution and sketch. In *CVPR*, pages 593–600, 2011.

[19] Z. Song, B. Ni, D. Guo, T. Sim, and S. Yan. Learning universal multi-view age estimator by video contexts. In *ICCV*, 2011.

[20] V. N. Vapnik. *Statistical Learning Theory*. John Wiley, New York, 1998.

[21] H. Wold. Path models with latent variables: The nipals approach. In H. M. Blalock and et al, editors, *Quantitative Sociology: Internnational perspectives on mathematical and statistical model building*, pages 307–357. Academic Press, 1975.

[22] H. Wold. Partial least squares. In S. Kotz and N. Johnson, editors, *Encyclopedia of Statistical Sciences*, volume 6, pages 581–591. Wiley, New York, 1985.

[23] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, and S. Lin. Graph embedding and extensions: A general framework for dimensionality reduction. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29:40–51, 2007.

[24] S. Yan, X. Zhou, M. Liu, M. Hasegawa-Johnson, and T. Huang. Regression from patch-kernel. In *IEEE conf. on CVPR*, 2008.