A Study on Cross-Population Age Estimation

Guodong Guo*
LCSEE, West Virginia University
guodong.guo@mail.wvu.edu

Chao Zhang
LCSEE, West Virginia University
cazhang@mix.wvu.edu

Abstract

We study the problem of cross-population age estimation. Human aging is determined by the genes and influenced by many factors. Different populations, e.g., males and females, Caucasian and Asian, may age differently. Previous research has discovered the aging difference among different populations, and reported large errors in age estimation when crossing gender and/or ethnicity. In this paper we propose novel methods for cross-population age estimation with a good performance. The proposed methods are based on projecting the different aging patterns into a common space where the aging patterns can be correlated even though they come from different populations. The projections are also discriminative between age classes due to the integration of the classical discriminant analysis technique. Further, we study the amount of data needed in the target population to learn a cross-population age estimator. Finally, we study the feasibility of multi-source cross-population age estimation. Experiments are conducted on a large database of more than 21,000 face images selected from the MORPH. Our studies are valuable to significantly reduce the burden of training data collection for age estimation on a new population, utilizing existing aging patterns even from different populations.

1. Introduction

Human age estimation is an active research topic in computer vision and pattern recognition in recent years [15] [3] [27] [26] [31] [2] [23], because of many potential applications [6] [21], e.g., age-specific human-computer interaction [9], and business intelligence [24]. However, age estimation is very challenging, especially on a large database with heterogeneous populations [11].

In previous research on age estimation using the large Yamaha Gender and Age (YGA) database of 8,000 face images [7, 29, 10, 14], it has been performed on the females and males, separately. Guo et al. [13] studied the problem

*This work was partially supported by an NSF CITeR grant.

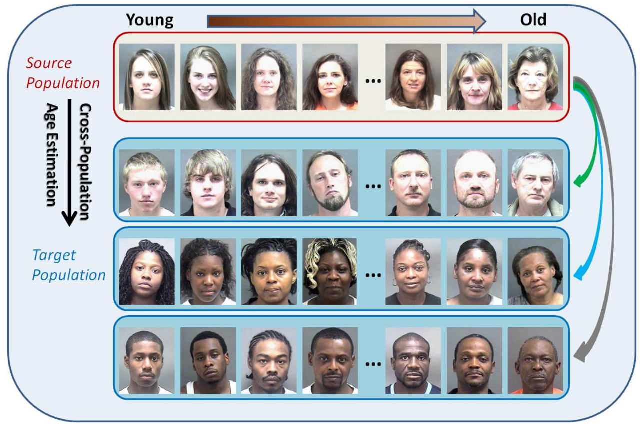

Figure 1. Illustrate the problem of *cross-population* age estimation. Here cross-population includes across different gender and/or ethnicity. Because of the gene differences and other factors, different populations may age differently.

quantitatively and showed the significant error increases when males and females are mixed. To deal with the influence of gender on age estimation, they proposed to recognize gender first and then do age estimation for each gender separately [13]. Ni et al. [19] performed cross-database age estimation using faces from the Internet as training data, and tested age estimation performance on some aging databases, such as the FG-NET [5] and MORPH [22]. However, the reported mean absolute errors (MAE) are very high. For example, the MAE is 8.60 years on MORPH database, and 9.49 years on FG-NET. Guo and Mu [11] studied the influence of gender and ethnicity on age estimation on the large MORPH database. They showed that cross either race or gender or both can cause very large errors in age estimation. To deal with the problem, they proposed to recognize the gender and ethnicity group first and then perform age estimation within each group. Another way to deal with the influence of gender and ethnicity is to estimate age, gender, and ethnicity together in one step [12].

So, previous research, e.g., [12, 11, 19, 13, 14, 10, 7, 29], has shown the aging difference among different populations in one way or another. Some approaches have been proposed to reduce the influence of population heterogeneity on age estimation [13, 19, 11, 12], however, there is no study yet to clearly address the problem of cross-population age estimation with a good result, e.g., a small MAE.

Considering the great difficulty in collecting aging image databases, it will have a remarkable value to develop an age estimator to work on a new population, utilizing the existing aging patterns in other different populations. Thus we propose to study a new problem called *cross-population age estimation*, as illustrated in Figure 1.

In this paper, we develop a method to address the novel problem of cross-population age estimation. The basic idea can be illustrated in Figure 2. Our method is based on projecting the different aging patterns into a "common" space where the aging patterns can be correlated even though they come from different populations. The projections are also discriminative between age classes since they integrate the classical discriminant analysis technique [18]. The method can also be used to study the amount of data needed in the target population to learn a good cross-population age estimator. We also propose to extend our method to multi-source cross-population age estimation.

In the following, we introduce our methods for cross-population age estimation in Section 2. Comprehensive experiments are conducted in Section 3, and finally, some discussions and conclusions are given.

2. Cross-Population Age Estimation

In this section, we present our method for cross-population age estimation. We first introduce the classical discriminant analysis with a slightly different formulation, which is helpful for us to develop our method to deal with cross-population aging pattern projections. Then we present our method. We also extend our method to deal with multi-source cross-population age estimation.

2.1. Linear Discriminant Analysis

Suppose we have a set of l samples $x_1, x_2, \cdots, x_l, \in R^d$, belonging to c classes. Let μ be the total sample mean vector, l_k the number of samples in the k-th class, and $x_i^{(k)}$ the $i-$th sample in the k-th class. The objective function of the classical linear discriminant analysis (LDA) [8] is $\hat{w} = \arg\max_w \frac{w^T S_b w}{w^T S_w w}$, where S_b and S_w are the between-class scatter matrix and within-class matrix, respectively. $S_b = \sum_{k=1}^{c} l_k(\mu^{(k)} - \mu)(\mu^{(k)} - \mu)^T$, $S_w = \sum_{k=1}^{c}\left(\sum_{i=1}^{l_k}(x_i^{(k)} - \mu^{(k)})(x_i^{(k)} - \mu^{(k)})^T\right)$.

Define the total scatter matrix $S_t = \sum_{i=1}^{l}(x_i - \mu)(x_i - \mu)^T$, then we have $S_t = S_b + S_w$ [8]. The traditional objective function of the LDA is equivalent to

$$\hat{w} = \arg\max_w \frac{w^T S_b w}{w^T S_t w}, \tag{1}$$

according to [8]. By normalizing the training samples to have a zero mean vector, i.e., subtracting the mean vector from all samples, we can have

$$\begin{aligned} S_b &= \sum_{k=1}^{c} l_k (\mu^{(k)})(\mu^{(k)})^T \\ &= \sum_{k=1}^{c} l_k \left(\frac{1}{l_k}\sum_{i=1}^{l_k} x_i^{(k)}\right)\left(\frac{1}{l_k}\sum_{i=1}^{l_k} x_i^{(k)}\right)^T \\ &= \sum_{k=1}^{c} X^{(k)} W^{(k)} (X^{(k)})^T = X W_{l\times l} X^T, \end{aligned} \tag{2}$$

where $W^{(k)}$ is a $l_k \times l_k$ matrix with all the elements equal to $1/l_k$, $X^{(k)} = [x_1^{(k)}, \cdots, x_{l_k}^{(k)}]$ denotes the data matrix of the k-th class, $X = [X^{(1)}, \cdots, X^{(c)}] \in R^{d\times l}$ and $W_{l\times l}$ is a $l \times l$ matrix:

$$W_{l\times l} = \begin{bmatrix} W^{(1)} & 0 & \cdots & 0 \\ 0 & W^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & W^{(c)} \end{bmatrix}. \tag{3}$$

Then the objective function of the traditional LDA can be re-formulated as

$$\hat{w} = \arg\max_w \frac{w^T X W_{l\times l} X^T w}{w^T X X^T w}, \tag{4}$$

which was first proposed by He et al. in [16]. This formulation is helpful to develop our cross-population discriminant analysis (CpDA). The solution of (4) could be found by solving the generalized eigenvalue problem as below:

$$X W_{l\times l} X^T w = \lambda X X^T w. \tag{5}$$

2.2. Cross-Population Discriminant Analysis

Given the discriminant analysis as formulated in (4), we will develop a method for our cross-population age estimation. It has been shown in [11] that the distributions of aging patterns are different in different populations, and the age estimation from one population to another can result in large errors. To deal with the problem of cross-population age estimation, we develop a method, called cross-population discriminant analysis (CpDA).

Suppose we have aging patterns from two populations: P for the source population, and Q for the target. Due to the distribution discrepancy of the aging patterns in P and Q, the optimal discriminative projections for the patterns in P are different from the learned projections for Q, when the traditional discriminant analysis is applied to them independently. So we need to consider aging patterns in both populations together to derive a common space for projections.

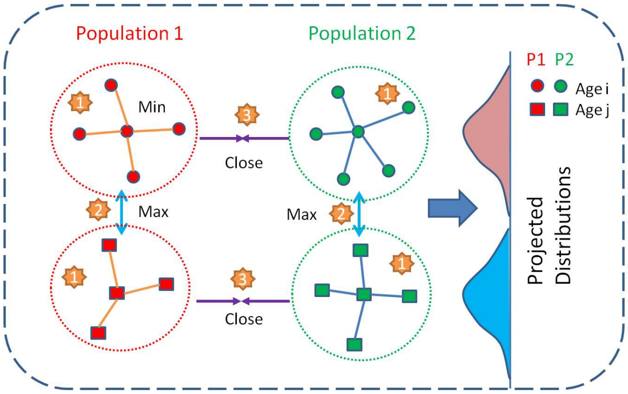

Figure 2. Illustrate the key idea of cross-population discriminant analysis (CpDA). Three measures are considered in the proposed method: (1) within-class variation, (2) between-class scatter, and (3) between-population correlation. After the mapping, the aging patterns from different populations are projected into a common space where similar or closer distributions could be resulted in.

Let $S_{P,b}$ and $S_{Q,b}$ be the between-class scatter matrix of the aging patterns in populations P and Q, respectively, and $S_{P,t}$ and $S_{Q,t}$ be the total scatter matrix for P and Q, separately. The objective function with aging patterns in two populations can be formulated as,

$$\hat{w} = \arg\max_{w} \frac{w^T(S_{P,b} + S_{Q,b})w}{w^T(S_{P,t} + S_{Q,t})w}. \tag{6}$$

However, this is a very naive formulation, since it just considers the two populations independently and ignores the interaction between the two populations.

Our idea for cross-population discriminant analysis is illustrated in Figure 2. We hope to find optimal projection directions to map the aging patterns from two different populations as close as possible, in addition to making them maximally separated among different age classes. Towards this goal, we introduce a term that measures the "correlation" or "closeness" between the aging patterns in two different populations,

$$\phi_{P,Q} = \sum_{k=1}^{c} (\theta_P^{(k)})(\theta_Q^{(k)})^T, \tag{7}$$

where $\theta_P^{(k)}$ is the sample mean vector in the k-th class of the aging patterns in population P after projections, say w, to be solved optimally, and $\theta_Q^{(k)}$ is the sample mean vector in the k-th class of the aging patterns in population Q after the same projections.

Let $l_P^{(k)}$ be the number of samples in the k-th age class in population P, and $l_Q^{(k)}$ be the number of samples in the k-th age class in population Q. Note that the numbers $l_P^{(k)}$ and $l_Q^{(k)}$ can be very different (see our experiments). Then Eqn. (7) can be computed as

$$\begin{aligned} \phi_{P,Q} &= \sum_{k=1}^{c} \left(\frac{1}{l_P^{(k)}} \sum_{i=1}^{l_P^{(k)}} y_{P,i}^{(k)} \right) \left(\frac{1}{l_Q^{(k)}} \sum_{j=1}^{l_Q^{(k)}} y_{Q,j}^{(k)} \right)^T \\ &= \sum_{k=1}^{c} \frac{1}{l_P^{(k)} l_Q^{(k)}} Y_P^{(k)} (Y_Q^{(k)})^T = \sum_{k=1}^{c} Y_P^{(k)} D^{(k)} Y_Q^{(k)} \end{aligned} \tag{8}$$

where $D^{(k)}$ is a $l_P^{(k)} \times l_Q^{(k)}$ matrix with all the elements equal to $1/l_P^{(k)} l_Q^{(k)}$ and $Y_P^{(k)} = \left[y_{P,1}^{(k)}, \cdots, y_{P,l_P^{(k)}}^{(k)} \right]$ denote the transformed data matrix of the aging patterns in the k-th class. Let the projection be w applied to the original aging patterns, i.e., $y^{(k)} = w^T x^{(k)}$, we have

$$\phi_{P,Q} = \sum_{k=1}^{c} w^T X_P^{(k)} D^{(k)} (X_Q^{(k)})^T w = w^T X_P D_{l_P \times l_Q} X_Q^T w, \tag{9}$$

where l_P is the number of total samples in population P, l_Q is the number of total samples in population Q. $D_{l_P \times l_Q}$ is a $l_P \times l_Q$ matrix with

$$D_{l_P \times l_Q} = \begin{bmatrix} D^{(1)} & 0 & \cdots & 0 \\ 0 & D^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D^{(c)} \end{bmatrix} \tag{10}$$

Let the matrix $S_d = X_P D_{l_P \times l_Q} (X_Q)^T$, we have $\phi_{P,Q} = w^T S_d w$. Then, the objective function of our cross-population discriminant analysis is:

$$\begin{aligned} \hat{w} &= \arg\max_{w} \frac{w^T(S_{P,b} + S_{Q,b})w + \beta\phi_{P,Q}}{w^T(S_{P,t} + S_{Q,t})w} \\ &= \arg\max_{w} \frac{w^T(S_{P,b} + S_{Q,b} + \beta S_d)w}{w^T(S_{P,t} + S_{Q,t})w}, \end{aligned} \tag{11}$$

where the item $\phi_{P,Q}$ can be considered as a regularizer which characterizes the correlation or closeness between the aging patterns in two different populations, β is to balance the between-class scatter and inter-population correlation. The optimal $w's$ are the eigenvectors corresponding to the generalized eigenvalue problem,

$$(S_{P,b} + S_{Q,b} + \beta S_d)w = \lambda(S_{P,t} + S_{Q,t})w. \tag{12}$$

Based on the projections, the aging patterns are expected to be transformed so that the aging differences between two different populations can be minimized and the aging patterns in different populations can be "pulled" towards similar distributions. Then the aging functions can be learned on the transformed aging patterns with classifiers or regression. In our work, we use the support vector machines (SVMs) [28] for aging function learning.

2.3. Multi-Source Cross-Population Discriminant Analysis

So far, we have addressed the cross-population discriminant analysis for age estimation with two populations: P as the source population and Q as the target. We are also interested in the case where more than one source population is available for learning an age estimator. Researchers may collect aging databases from different populations in different countries, thus it is valuable to study multi-source age estimation.

Here we will extend the formulation in (12) to a multi-source cross-population discriminant analysis.

The key idea for multi-source cross-population analysis is to maximize the correlation between any two populations, including all sources to the target as well as among all source populations.

Denote the target population as Q with a training set of aging patterns, and the n source populations as $P_1, \cdots, P_n$. By extending the single source cross-population discriminant analysis (CpDA), we can get the multi-source cross-population discriminant analysis as follows,

$$S_B = \sum_{r=1}^{n} S_{P_r,b} + S_{Q,b},\ S_T = \sum_{r=1}^{n} S_{P_r,t} + S_{Q,t}, \quad (13)$$

$$S_D = \sum_{r=1}^{n} S_d(P_r, Q) + \sum_{s=1}^{n} \sum_{t=s+1}^{n} S_d(P_s, P_t). \quad (14)$$

Then the objective function of multiple-source cross-population discriminant analysis is given by $\hat{w} = \arg\max_w \frac{w^T(S_B+\beta S_D)w}{w^T S_T w}$, which can be solved by the generalized eigenvalue problem, similar to (12). So a nice property here is that our formulation can be easily extended to multi-source age estimation.

3. Experiments

We evaluate the proposed methods experimentally on the public available MORPH database [22]. The FG-NET database [5] is also public available, but it is too small to get a statistically meaningful result. The MORPH database is probably the only large database that contains populations of different races, to the best of our knowledge.

We introduce the database first and then present the experimental results. Our major focus is on *cross-population* age estimation.

3.1. The Database

The MORPH [22] is a large database containing two sections, I and II. Since MORPH-I is too small (1,690 face images), we use MORPH-II for our study. The MORPH-II has about 55,000 face images, containing faces of various races. However, the distribution of gender and ethnicity is very unbalanced. For instance, it has about 77% Black faces, 19% White, and 4% other races, e.g., Hispanic, Asian, and Indian. There are also more males than females.

To avoid the influence of unbalanced distributions, we selected data from MORPH-II with almost all White faces and a matched number of Black faces to build a relatively balanced database for our study. Actually, this unbalance problem was noticed in [11], and a database was selected for their study. We follow the similar consideration as in [11], and assembled a database of 21,060 face images. Specifically, there are 2,570 White Female (WF), 7,960 White Male (WM), 2,570 Black Female (BF), and 7,960 Black Male (BM) face images. The age range is from 16 to 67 years in the assembled database, the same as the whole MORPH-II.

The selected data within each population, e.g., WF, is randomly divided into two sets with an equal size. Then one set can be used for training, while the other for testing (of different crossings). This way, we can also perform age estimation within the same group or population if needed, and there is no overlap between any training and testing data. We use the biologically inspired features (BIF) [14] for facial image representation in our study. The face images are aligned with detected eye centers, and cropped and normalized into 60x60. The parameter β is set as 0.5 in the CpDA method in all our experiments.

3.2. Results of Cross-Population Age Estimation

The age estimation results using the proposed method are shown in Table 1. The age estimation performance is usually measured by the mean absolute errors (MAE), which is defined as the average of the absolute errors between the estimated ages and the ground truth ages.

The previous approach [11] is a direct age estimation where large errors obtained in cross gender and/or race. Specifically, the aging patterns in one population are used for training, while a different population is used for testing. There is no learning of the relations between different populations. Since [11] is the only work related to cross-population age estimation, we list their results in Table 1 to show if a learning-based method can help or not for cross-population age estimation.

From Table 1, we can see that our learning-based method can significantly reduce the errors in each cross case, comparing to the results without learning [11]. In row 1, the Black Female (BF) is used as the source population, while others as the target populations, including White Female (WF), Black Male (BM), and White Male (WM). The direct approach [11] to cross-population age estimation, BF→WF, has a MAE of 9.15 years. Our CpDA method can reduce the MAE to 6.41 years, with an error reduction rate (ERR) of 29.9%. In the next two lines, the BM and WM are used as the target population, separately. The estimation errors can have significant reductions.

Other cases of cross-populations age estimation are shown in rows 2∼4, using our proposed learning-based method. We can observe from Table 1 that the MAE can be reduced significantly for all cases of cross-population age estimation.

From this experiment, we show that a learning-based method can be developed to deal with cross-population age estimation. This result can make an impact in practice, since it is usually very difficult to collect aging faces with accurate age labels for each new population.

Table 1. Cross-population age estimation using different methods. "BF" denotes Black Female, "WF" for White Female, "BM" for Black Male, "WM" for White Male. The method in [11] is a direct age estimation without learning, "CpDA" is the proposed learning-based method. "MAE" is the mean absolute error in years.

Train	Test	Method and Performance	
		W/O Learning [11]	CpDA (New)
		MAE(yrs.)	MAE(yrs.)
BF	WF	9.15	6.41
	BM	8.40	6.13
	WM	8.79	6.23
WF	BF	8.67	6.54
	WM	7.72	5.57
	BM	10.62	7.54
BM	WM	6.86	5.10
	BF	10.58	7.73
	WF	12.81	8.73
WM	BM	7.05	5.35
	WF	9.13	6.70
	BF	9.54	7.67

3.3. Percentage of Data in the Target Population

In this experiment, we study the amount of training data needed in the target population to learn a cross-population age estimator with good performance. The purpose of this study is to explore if a smaller number of aging faces in a new population can be sufficient for learning. For this study, we fix the number of training examples in the source population, but change the number of training examples in the target population in percentage, ranging from 10% to 100%. We use percentages to control the number of examples, since there are different numbers of images in different populations. The testing data are fixed in the target population. Remember that we have two sets for each population. One is for training while the other for testing (in different crossings). The percentage changes are only in the training set. The comprehensive experiments are performed using the proposed method CpDA in each case. The SVMs are used for aging function learning after feature transformation with our method.

Table 2. MAEs (in years) based on some representative transfer learning methods, compared to our method. The study case is BF→WF with different amount of target data. The "A-SVM" method uses linear (LIN) or RBF kernel, but the errors are too high, even higher than the results in [11]. The "A-MKL" using multiple kernels, is extremely slow in our problem. Its estimation errors are also higher than our method.

Method	10%	20%	30%	40%	50%
A-SVM(LIN)	17.80	19.58	24.96	25.08	24.53
A-SVM(RBF)	20.89	21.19	23.02	24.14	23.95
A-MKL	7.56	7.54	7.35	7.34	6.89
CpDA (ours)	6.90	6.81	6.76	6.46	6.46

To show the results visually, we draw curves corresponding to three average cases of cross-population age estimation. The curves are the MAEs w.r.t. the percentage of data in the target population for training, as shown in Fig. 3. Each curve is an average of several results. For example, the curve of "From Black to White" based on the CpDA is the average of BF→WF and BM→WM. We can observe that our method, i.e., the CpDA, performs quite well. We also observe that a small amount of training data in the target population is sufficient to learn the cross-population age estimator with a good performance. For instance, about 20% of the training data in the target population can be enough to have a MAE which is within one year difference from the 100% training data. About 30% of the training data in target can be sufficient to have a MAE within 0.5 years difference from the 100% training data. This finding is very useful in practice, since we may just need to collect a small amount of training data from a new population but can still do age estimation by utilizing the existing aging patterns in another population. In other words, we can take advantage of the aging patterns in another population rather than discarding them.

3.4. Comparison with Other Methods

Our cross-population age estimation might be considered as a transfer learning problem [20]. However, age estimation is much more challenging than a regular object recognition problem. One aspect is that the separation between different ages is not so distinct as between different objects or visual events. Another is that there are usually many age classes or labels (each year), much more than the number of classes in many current vision problems using transfer learning techniques, e.g., [30] [4]. To show the challenges in our problem, we apply some popular transfer learning methods in computer vision for comparisons. One is the adaptive SVM or A-SVM [30], and the other is adaptive multiple kernel learning (A-MKL) [4]. Both methods work well on other vision problems, e.g., visual concept detection or event recognition across different domains.

The comparisons are shown in Table 2 using different

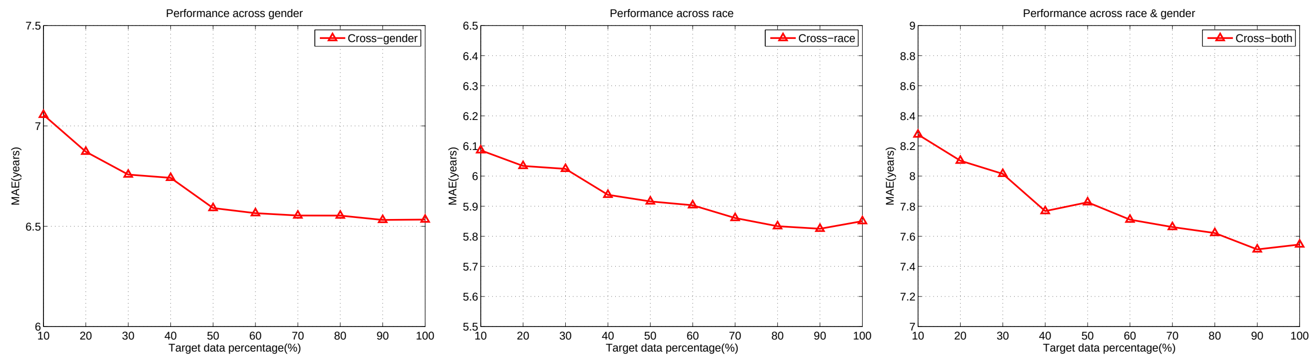

Figure 3. The MAEs of age estimation results w.r.t. the percentage of the training data in the target population. The MAEs reduce when more training data are used in the target population. About 30% of the target data is sufficient to make the MAE reduce to less than about 0.5 year difference from the result using 100% training data in the target population.

percentages of the target data for cross-population learning. From the table, we find that the A-SVM cannot work well on our cross-population age estimation. The MAEs of the A-SVM method are much higher than our methods, no matter what kernel is used: linear or RBF. The errors are even higher than the direct age estimation without learning between populations [11] (see Table 1). For the A-MKL method, the MAEs are larger than our proposed method in each case. More importantly, we found that the A-MKL method is extremely slow in our problem. It takes more than one day to obtain one item result in Table 2, in our 64-bit computer with i7(3.4G) CPU processor and 12G RAM. While it only takes about three hours for our method to get all results. Note that we used the same computer with the same data for comparisons between different methods. So our method performs much better than both the A-SVM and A-MKL methods. The comparisons also indirectly indicate that the cross-population age estimation is a very challenging problem.

3.5. Multi-Source Cross-Population Age Estimation

In the last experiment, we explore another new problem called *multi-source cross-population* age estimation. The idea is to use more than one source population for learning. In this way, we can utilize more training data from different populations in order to have a good age estimator for a new population. We verify the combination of two populations as the sources. For example, BF and WF are two sources for the target population BM. The multi-source cross-population age estimation results are shown in Table 3. The target data have different percentages for training, e.g., from 10% to 50%, while two source populations are used for training with all available training data. The MAE results in Table 3 are compared to the single-source cross-population age estimation results. For example, the MAE of BF+WF→BM is compared to either BF→BM or WF→BM. From Table 3, we can observe that in most cases, the combination of two source populations can reduce the estimation errors, especially for the larger error of the two corresponding single-source cross-population results. Specifically, among the 40 results, there are 22 cases (highlighted by the red color) where the combinations achieve a MAE less than either of the two single-source estimation result, and 18 cases where the combinations reduce the larger errors of the two single-source age estimation results. Based on this experiment, we can see that it is helpful to perform multi-source cross-population age estimation in that all available aging patterns can be utilized to improve the accuracies.

Table 3. Results of multi-source cross-population age estimation based on our method, in terms of different percentages of training data in the target population for training. Compared to the single-source results, the 40 results have 22 cases (highlighted by the red color) where the multi-source approach performs better than using each single source, 18 cases where the multi-source approach reduces the larger error of the two single-source results.

Train	Test	10%	20%	30%	40%	50%
BF+WF	BM	6.61	6.65	6.57	6.53	6.47
BF+WF	WM	5.86	5.88	5.83	5.59	5.70
WM+BM	WF	6.90	6.63	6.53	6.65	6.58
WM+BM	BF	6.47	6.57	6.59	6.43	6.40
BF+BM	WF	6.81	6.76	6.79	6.61	6.59
BF+BM	WM	5.33	5.22	5.31	5.29	5.23
WF+WM	BF	6.73	6.67	6.53	6.46	6.32
WF+WM	BM	6.07	5.93	5.97	5.94	5.96

4. Discussions

Linear discriminant analysis (LDA) is a classical method for pattern recognition [18]. He et al. [16] presented another formulation of the LDA which is the basis for the development of our cross-population discriminant analysis method. There are other extensions of the LDA, such as the semi-supervised discriminant analysis (SDA) [1], multi-view discriminant analysis (MDA) [17, 25], and so on. Those problems are different from our cross-population age estimation. For example, the SDA focuses on the utilization of unlabeled data that are assumed to have the same distribution as the labeled data; the MDA deals with data that are assumed to belong to the same object or subject at different views. Our cross-population age estimation cannot be considered as a multiview problem since there does not exist different views of the same subject in different populations. So our problem is different from either the SDA or MDA, and the assumptions in SDA or MDA cannot be satisfied in our case.

On the other hand, it seems that our problem might be considered as a transfer learning problem. Thus we evaluated some recently developed transfer learning techniques, e.g., [30] [4], for our problem. From the experimental results, those methods cannot work well on our problem, either because of large errors or extremely slow, although these methods have shown great successes in other vision tasks. These results indirectly indicate that the cross-population age estimation is a challenging problem. Although we have got some good results, more research efforts are needed to further improve the performance.

5. Concluding Remarks

We have studied a novel problem called *cross-population age estimation*. A method has been proposed to solve the challenging problem based on the discriminant analysis of aging patterns within and across populations. The proposed method has been evaluated on a large database with more than 21,000 face images, and the experimental results have shown that an appropriate learning-based method can be developed to deal with cross-population age estimation. We have also studied the amount of training data needed in the target population, and found that about 30% of training data in the target population is sufficient to achieve a good performance. This result can have a great value in practice, since it can largely reduce the difficulty in aging face image collection for a new population. Finally, we have shown that a multi-source cross-population age estimator can be developed that can effectively utilize exiting aging patterns in several different populations.

References

[1] D. Cai, X. He, and J. Han. Semi-supervised discriminant analysis. In *IEEE Int'l Conf. on Computer Vision*, pages 1–7, 2007.

[2] K. Chang, C. Chen, and Y. Hung. A ranking approach for human ages estimation based on face images. In *ICPR*, pages 3396–3399, 2010.

[3] K. Chen, S. Gong, T. Xiang, Q. Mary, and C. C. Loy. Cumulative attribute space for age and crowd density estimation. In *CVPR*, 2013.

[4] L. Duan, D. Xu, I. Tsang, and J. Luo. Visual event recognition in videos by learning from web data. *TPAMI*, 34(9):1667–1680, 2012.

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

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

[7] Y. Fu and T. S. Huang. Human age estimation with regression on discriminative aging manifold. *IEEE TMM*, 10(4):578–584, 2008.

[8] K. Fukunaga. *Introduction to statistical pattern recognition*. Academic Press, 2nd edition edition, 1990.

[9] X. Geng, Z.-H. Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. *TPAMI*, 29(12):2234–2240, 2007.

[10] 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.

[11] G.-D. Guo and G. Mu. Human age estimation: what is the influence across race and gender? In *IEEE Int'l Workshop on Analysis and Modeling of Faces and Gestures*, 2010.

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

[13] 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 Int'l Conf. on Computer Vision*, pages 1986–1991, 2009.

[14] G.-D. Guo, G. Mu, Y. Fu, and T. S. Huang. Human age estimation using bio-inspired features. In *IEEE CVPR*, pages 112–119, 2009.

[15] H. Han, C. Otto, and A. K. Jain. Age estimation from face images: human vs. machine performance. In *Int'l Conf. on Biometrics*, 2013.

[16] X. He, S. Yan, Y. Hu, P. Niyogi, and H.-J. Zhang. Face recognition using laplacianfaces. *IEEE TPAMI*, 27(3):328–340, 2005.

[17] M. Kan, S. Shan, H. Zhang, S. Lao, and X. Chen. Multi-view discriminant analysis. In *ECCV*, pages 808–821. 2012.

[18] G. McLachlan. *Discriminant analysis and statistical pattern recognition*, volume 544. Wiley-Interscience, 2004.

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

[20] S. Pan and Q. Yang. A survey on transfer learning. *IEEE Trans. on Knowledge and Data Engineering*, 22(10):1345–1359, 2010.

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

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

[23] C. Shan. Learning local features for age estimation on real-life faces. In *ACM W. on Multimodal Pervasive Video Analysis*, 2010.

[24] C. Shan, F. Porikli, T. Xiang, and S. Gong. *Video Analytics for Business Intelligence*, volume 409. Springer, 2012.

[25] A. Sharma, A. Kumar, H. Daume, and D. W. Jacobs. Generalized multiview analysis: A discriminative latent space. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 2160–2167, 2012.

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

[27] P. Thukral, K. Mitra, and R. Chellappa. A hierarchical approach for human age estimation. In *ICASSP*, pages 1529–1532, 2012.

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

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

[30] J. Yang, R. Yan, and A. Hauptmann. Cross-domain video concept detection using adaptive svms. In *Int'l conf. Multimedia*, pages 188–197, 2007.

[31] P. Yang, L. Zhong, and D. Metaxas. Ranking model for facial age estimation. In *ICPR*, pages 3404–3407, 2010.