A Study on Automatic Age Estimation using a Large Database

Guodong Guo
WVU

Guowang Mu
NCCU

Yun Fu
BBN Technologies

Charles Dyer
UW-Madison

Thomas Huang
UIUC

Abstract

In this paper we study some problems related to human age estimation using a large database. First, we study the influence of gender on age estimation based on face representations that combine biologically-inspired features with manifold learning techniques. Second, we study age estimation using smaller gender and age groups rather than on all ages. Significant error reductions are observed in both cases. Based on these results, we designed three frameworks for automatic age estimation that exhibit high performance. Unlike previous methods that require manual separation of males and females prior to age estimation, our work is the first to estimate age automatically on a large database. Furthermore, a data fusion approach is proposed using one of the frameworks, which gives an age estimation error more than 40% smaller than previous methods.

1. Introduction

Human age estimation has recently become an active research topic in computer vision because of the demands of real-world applications such as electronic customer relationship management (ECRM) [2], security control and surveillance monitoring [10, 16, 20], biometrics [21, 19], and entertainment.

Computationally, an age estimation system usually consists of two modules: image representation and age estimation. Age image representations include the anthropometric model [15, 20], wrinkle model [14], active appearance model (AAM) [5], AGing pattErn Subspace (AGES) [9, 8], age manifold learned from raw images [7, 6], local binary pattern features [34], and parts [25] or patch-based appearance model [33, 29]. Given a representation, age estimation can be viewed as a multi-class classification problem [16, 26, 10] or a regression problem [7, 31, 30, 6, 35] or a hybrid of the two [10, 11, 12].

The aging process is influenced by many factors. For example, males and females may have different face aging patterns [10]. So far, however, no research has studied the influence of gender on age estimation. We will study this problem in this paper. Furthermore, we want to study age estimation performance on separated age groups rather than on all ages. The results are related to designing an *automatic* age estimation system without user's input of gender.

All our studies have to be performed on a large database because it is hard to discover the problems mentioned above using a small database. We first introduce the database used for our study in Section 2. Face representations are described in Section 3. Then the studies and results are presented in Section 4. After that, we introduce several frameworks for automatic age estimation in Section 5.

2. The Database

The Yamaha gender and age (YGA) database was adopted in our research. It is a large face database collected recently, containing 8,000 face images captured outdoors for people in the age range from 0 to 93 years. The database has been used before for human age estimation [7, 31, 30, 11, 6, 10, 33, 12]. But in all of those approaches, males and females were *manually separated* before age estimation, which is not practical for a completely automatic age estimation system.

3. Face Representations

We explore the combinations of biologically inspired features (BIF) with manifold learning techniques for face representations, which will be used in our study of problems related to age estimation. Given the face representations, support vector machines (SVMs) [27] were used for age estimation, i.e., each age is used as one class label. Multi-class SVMs based on pairwise comparisons have been used successfully for age estimation previously [10]. In this paper we also use SVMs, with a variety of face representations.

3.1. Biologically-Inspired Features

Visual processing in the cortex is modeled as a hierarchy of increasingly sophisticated representations. Riesenhuber and Poggio [22] proposed a new set of features derived from a feed-forward model of the primate visual object recognition pathway. The model contains alternating layers of simple (S) and complex (C) cell units. The bio-inspired features

(BIFs) have been investigated for object category recognition [24] [23] [18] and face recognition [17]. Very recently, BIFs were investigated for age estimation [13] and showed good performance without using learned prototypes as in [24, 23, 18]. Here, BIFs are used as the base features in our study. We found that age estimation performance can be improved significantly when manifold learning uses BIF features.

3.2. Manifold Learning

Here we investigate two recently proposed manifold learning methods, called marginal fisher analysis (MFA) [32] and locality sensitive discriminant analysis (LSDA) [4], for the age estimation problem. An earlier method, called orthogonal locality preserving projections (OLPP) [3], has been applied to age estimation in [7, 11, 10]. We include the OLPP method in our study for comparison, and found that when OLPP is used with the BIF rather than raw images as in [7, 11, 10], a significant gain in performance can be obtained. Principal component analysis (PCA) [28] is also included for comparison.

3.3. Our Representations

We investigate combinations of biologically-inspired features and manifold learning as the face representations for age estimation. Unlike unsupervised PCA, the methods that we investigate here, such as LSDA and MFA, can incorporate label information during learning. Manifold learning or subspace analysis is an active research topic in computer vision recently. Usually, manifold learning methods are applied to raw images directly [3, 32, 4, 7, 11, 10, 6]. However, there might be a problem in applying subspace analysis methods to raw images directly because misalignment of images can cause problems in learning a discriminative, low-dimensional manifold. In other words, subspace analysis methods are often sensitive to image misalignment. In our age estimation problem, the face images have a large span of ages, e.g., from 0 to 93 years, so it is very hard to align all these images well, considering the large facial shape changes from young to senior people. How to deal with this problem? Using biologically-inspired features is one possible solution. In deriving the BIF, the "MAX" operation is used to obtain the C_1 features from S_1, which can endure small translations, rotations, and scale changes [22, 23].

We believe that the combination of BIF and manifold learning is promising in terms of robustness (i.e., not sensitive to image misalignment) and discriminative power (local embedding by supervised learning with label information).

4. Study and Results

Based on the face representations described in the previous section, we study two issues in human age estimation: the effects of gender and the effects of gender and age group on age estimation. These issues are very related to designing *automatic* age estimation methods with high performance, which will be presented in our next study.

4.1. Gender Unknown vs. Known in Age Estimation

Our first study is: *How much can age estimation performance be affected by gender?* For this investigation, we compare age estimation errors under two situations: case I, age estimation is performed on all faces without discriminating between males and females; and case II, age estimation is performed for males and females separately, with the assumption that gender is known.

To our best knowledge, no previous work has studied the issue of how gender affects age estimation. We compare age estimation errors in cases I and II using a variety of facial representations. The basic representation is to learn the age manifold from raw face images using the OLPP method, which has shown good performance for age estimation [7, 11, 10, 6]. It was also shown that the OLPP method is much better than PCA in learning the age manifold [7]. In addition to using the OLPP method for males and females separately as in [7, 11, 10, 6], we also apply OLPP to the whole training set without separating males and females. Linear SVMs were used with pairwise comparisons for age estimation. The performance of age estimation is usually measured by the mean absolute error (MAE) [16, 9], which is defined as the average of the absolute errors between the estimated ages and the ground truth ages. As shown in columns I and II of the first row in Table 1, the MAE is 7.04 years when the OLPP was applied to all training face images, while the MAEs are 5.24 and 5.69 years when it was applied to females and males separately. The average of the MAEs over females and males is 5.47 years, which is reduced from 7.04 years by 22.3%. This large reduction of error shows that age estimation is affected by gender significantly.

Throughout this paper the parameters for manifold learning and SVM training were adjusted using a small tuning set, which was about 10% of the training data. Then manifold learning was executed again and the SVMs retrained on the whole training set based on the tuned parameters. All results are measured by a standard 4-fold cross validation.

The second representation for age estimation is to use biologically-inspired features with PCA for dimensionality reduction [13]. Again, the work in [13] only performed age estimation for females and males separately. The MAE of the "BIF+PCA" representation is reduced by 14.3%, from 4.56 years to 3.91 years (on average) when age estimation

Table 1. A study of age estimation in different situations: case I, age estimation without considering gender separation; case II, age estimation with known gender (males and females are manually separated); and case III, age estimation on separated gender and age groups (group information is provided manually). The mean absolute errors (MAEs) in years are reported for age estimation in each case.

Representations	I	II		III					
	Altogether	Gender Known		Gender & Age Groups Known					
		Female	Male	Female			Male		
				Young	Adult	Senior	Young	Adult	Senior
OLPP [7, 11, 10, 6]	7.04	5.24	5.69	1.18	3.88	2.71	1.46	3.61	2.26
		Average: 5.47		Average: 2.88					
		Reduced: 22.3%		Reduced: 59.1%					
		Female	Male	Female			Male		
				Young	Adult	Senior	Young	Adult	Senior
BIF + PCA [13]	4.56	4.32	3.50	1.07	3.53	2.40	1.15	2.89	1.55
		Average: 3.91		Average: 2.43					
		Reduced: 14.3%		Reduced: 46.7%					
		Female	Male	Female			Male		
				Young	Adult	Senior	Young	Adult	Senior
BIF + OLPP (New)	3.88	2.89	2.76	0.92	2.89	1.63	1.09	**1.91**	1.71
		Average: 2.83		Average: 1.90					
		Reduced: 27.1%		Reduced: 51.0%					
		Female	Male	Female			Male		
				Young	Adult	Senior	Young	Adult	Senior
BIF + LSDA (New)	3.62	2.87	2.58	0.83	**2.34**	1.70	0.82	2.17	**1.28**
		Average: 2.73		Average: 1.74					
		Reduced: 24.6%		Reduced: 51.9%					
		Female	Male	Female			Male		
				Young	Adult	Senior	Young	Adult	Senior
BIF + MFA (New)	3.17	2.61	2.64	**0.79**	2.42	**1.56**	**0.76**	2.09	1.34
		Average: 2.63		Average: 1.72					
		Reduced: 17.0%		Reduced: 45.7%					

was performed for females and males separately. By the way, the "BIF+PCA" representation gives much smaller age estimation errors than OLPP on raw images in both cases.

Next, we applied OLPP to the BIF features. Age estimation results are shown in row three in Table 1. In contrast to PCA, the OLPP method can help the BIF lower the MAE from 4.56 to 3.88 years in case I. The MAE on average over males and females (case II) is also lowered from 3.91 to 2.83 years. This demonstrates that OLPP is better than unsupervised PCA given the same biologically-inspired features.

Let us also look at the difference of MAEs between cases I and II using the "BIF+OLPP" representation. The MAE on average in case II is 2.83 years, while the MAE is 3.88 years in case I. From case I to II, the error reduction rate is 27.1%. This significant error reduction demonstrates again the great influence of gender on age estimation.

Next, we examine manifold learning on BIF using the LSDA method. In case I, the MAE is 3.62 years, while in case II, it is reduced to 2.73 years, with a reduction rate of 24.6%. In addition, the "BIF+LSDA" representation gives slightly lower errors than the "BIF+OLPP" in both cases.

Finally, we examine manifold learning with MFA on BIF. In case I, the MAE is 3.17 years. It is reduced to 2.63 years in case II with an error reduction rate of 17.0%. Among the three manifold learning methods applied to BIF, the "BIF+MFA" representation has the fewest errors.

In summary, the influence of gender on age estimation is significant. When age estimation is performed on females and males separately, the MAEs can be reduced from 14.3% to 27.1%, based on the five methods for face representation.

4.2. Gender and Age Groups

Our second study addresses the question: *How significantly can age estimation errors be reduced if the estimations are performed on smaller gender and age groups?* We call it case III in our study, where gender and age groups are known or provided by the user. Motivated by the physical difference of human aging in different growing stages [1], we would like to investigate whether a gain in performance can be obtained if age estimation is performed on age groups rather than on all ages. From a computational viewpoint, the problem of age estimation is simplified a lot when working on smaller age groups.

To verify this conjecture, we divided ages into three groups: young (0-19 years, 1,000 images), adult (20-60, 2,050), and senior (61-93, 950) for both males and females. Then age estimation is performed on each group independently. The MAEs are reported for each group for case III in Table 1. The average over all six groups is computed by a weighted average, considering the different number of images in the groups. All results shown in Table 1 are measured by a standard 4-fold cross validation.

Error reduction rates were computed by comparing cases I and III. Given the five representations, the error reduction rates were 59.1%, 46.7%, 51.0%, 51.9%, and 45.7%, respectively. The reductions are significant, much greater than the corresponding reductions from case I to II. This demonstrates that age estimation errors can be very low if the problem is simplified to work on groups containing a small range of ages. The results in case III can be viewed as *"lower bounds"* on age estimation errors for the five representations.

4.3. Comparison with Other Approaches

Before moving to the next study, this section describes some comparisons. All previous work [31, 30, 11, 10, 33, 12] did age estimation for females and males separately, similarly to case II in our study. In Table 2, all results are reported using four-fold cross validation. The results of the last three are based on our new representations. In comparison, the new representations give much lower errors than the result of RPK [33], which has an average MAE of 4.66 years. Even the "BIF+OLPP" representation lowers the MAE by 39.3%. The "BIF+MFA" reduced the MAE further to 2.63 years with a reduction rate of 43.6%. These comparisons show that our new representations are better than previous approaches for age estimation.

Table 2. MAEs (in years) on the YGA database based on new representations (the last three rows) and previous methods for females, males, and all, respectively.

Method	YGA:F	YGA:M	Average
OLPP+SVR [10]	7.00	7.47	7.24
OLPP+SVM [10]	5.55	5.52	5.54
LARR [10]	5.25	5.30	5.28
PFA [12]	5.11	5.12	5.12
RPK [33]	4.94	4.38	4.66
BIF+OLPP (New)	2.89	2.76	2.83
BIF+LSDA (New)	2.87	**2.58**	2.73
BIF+MFA (New)	**2.61**	2.64	**2.63**

5. Automatic Age Estimation

According to the results in Table 1, age estimation performance can be improved significantly if the gender is known (case II), and improved even more if the gender and age groups are also provided (case III). In reality, however, this information is usually unknown for a test image. To develop an age estimation system for practical use, the gender or gender and age groups have to be recognized before age estimation. The question is, what performance can be achieved for completely automatic age estimation? To our best knowledge, no previous research has addressed this problem using a large database.

We present several different approaches to automatic age estimation, based on the results in Table 1. The four different face representations: "BIF+PCA," "BIF+OLPP," "BIF+LSDA," and "BIF+MFA," are used in each approach. They are also compared to see the difference between them.

5.1. Preceding Age Estimation with Gender Classification

A straightforward approach to automatic age estimation is first to recognize the gender of a given face, then do age estimation corresponding to that gender. If the face is classified as female, its age is estimated using the age estimator learned for females. Otherwise, the estimator learned for males is used. We call this approach framework 1 (F1).

Table 3. Automatic age estimation: framework 1 (F1).

Representations	Gender Classification Accuracy	Age Estimation Error (MAE)
BIF+PCA	93.7%	4.10
BIF+OLPP	92.6%	3.21
BIF+LSDA	**94.5%**	**2.95**
BIF+MFA	91.4%	3.12

The results of F1 with different representations are shown in Table 3. Comparing Table 3 with Table 1, it can be seen that (1) automatic age estimation using F1 can have better results than the corresponding mixed age estimation for all four representations, although the gender classifications are not perfect; (2) the best representation is the "BIF+LSDA" for age estimation with F1, which has the highest gender classification accuracy among the four; (3) the "BIF+MFA" representation performs the worst with F1 for gender classification, thus raising the MAE, although it has the smallest MAE for age estimation with mixed gender; and (4) all representations using F1 have larger errors than when gender is known.

5.2. Preceding Age Estimation with Gender and Age Group Classification

Given the smallest age estimation errors in case III of our study, another framework can be designed for automatic age estimation. It classifies a face into one of the gender and age groups first, and then does age estimation on the determined group only. We call this framework 2 (F2). This approach is reasonable if the accuracy of group classification is high enough so that the estimation errors are very small, as shown in the study of case III in Section 4. The results of F2 are shown in Table 4, from which one can observe that (1) the group classification accuracy can be as high as 89.7% using the "BIF+LSDA" representation. The final MAE can be as low as 2.84 years, which is lower than the 2.95 years using F1. (2) For most of the representations, the performance using F2 is better than that with F1. The only exception is "BIF+MFA." Notice that the group classification accuracy is too low for "BIF+MFA," which causes its low performance using F2.

Table 4. Automatic age estimation: framework 2 (F2).

Representations	Gender and Age Group Classification Accuracy	Age Estimation Error (MAE)
BIF+PCA	87.8%	3.63
BIF+OLPP	86.7%	3.19
BIF+LSDA	**89.7%**	**2.84**
BIF+MFA	82.4%	3.76

5.3. Using Gender Only from Gender and Age Group Classification

An alternative use of the gender and age group classification results is to use the estimated group information to make a new decision on gender. Specifically, if a test face is classified into one of the three groups: young female, adult female, or senior female, it is considered female. On the contrary, if the test face is classified into one of the three groups: young male, adult male, or senior male, it is determined as a male. Is gender decided in this way better than direct classification (e.g., in F1)? Consider the results in Table 5. Compared with Table 3, the gender recognition accuracies are improved for all four representations. The highest accuracy is now 95.6% for the "BIF+LSDA" representation. Even for "BIF+MFA," gender classification accuracy is slightly improved from 91.4% to 92.2%. The MAEs are also reduced for most of the representations. The only exception is "BIF+MFA." Comparing Tables 5 and 4, the MAEs using F3 are larger than that with F2 for the"BIF+LSDA" and "BIF+PCA" representations. The main advantage of F3 is that it can improve the gender recognition accuracies over the direct two-class gender classification used in F1.

We want to mention that nonlinear SVMs with RBF kernel perform much better than linear SVMs for gender classification. This is also true for gender and age group classification. But for age estimation, linear SVMs are comparable with or even better than the kernel SVMs.

Table 5. Automatic age estimation: framework 3 (F3).

Representations	Gender from Age and Gender Group	Age Estimation Error (MAE)
BIF+PCA	94.9%	4.04
BIF+OLPP	94.1%	2.98
BIF+LSDA	**95.6%**	**2.87**
BIF+MFA	92.2%	3.16

5.4. A Data Fusion Approach

Given the three frameworks for automatic age estimation, which of the four face representations gives the best results? The tables show that the "BIF+LSDA" representation performs the best consistently. Notice that it has the highest accuracy for gender and age group classification. On the other hand, if we go back to case III in Table 1, the smallest MAEs for age estimation are obtained from different representations in different groups. For example, the "BIF+MFA" representation has the lowest MAEs in three groups: 0.79 years for the young female, 1.56 years for the senior female, and 0.76 years for the young male, while the "BIF+LSDA" representation gives the lowest MAEs in two groups: 2.34 years for the adult female and 1.28 years for the senior male, and the "BIF+OLPP" gives the lowest MAE of 1.91 years for the adult male. So a natural question is, can we combine the different representations to further improve age estimation performance? This is an idea of data fusion.

To develop an approach based on data fusion, we propose a simple method to utilize the results of case III in Table 1. The fusion method works using F2 because only F2 can take full advantage of the smallest MAEs in each of the six groups. It classifies a face into one of the six groups first and then estimates age within that group. In contrast to F2 described in Section 5.2, now different representations are used in different groups, guided by case III in Table 1. Only the "BIF+LSDA" representation is used for gender and age group classification. We label this data fusion approach using framework 2 as "F2.F" and show the results in Table 6. The MAE is 2.66 years, which is the lowest error that we achieve for automatic age estimation.

5.5. Comparison

We compare our automatic age estimation results with the method in [33], although the comparison is not fair to us because all our results in Section 5 were obtained using automatic age estimation while previous work used manual separation of males and females in testing. The purpose of the comparison is to summarize our results from Table 3 to Table 6, and show the improvements.

Table 6. Automatic age estimation: a data fusion approach using framework 2 (F2.F).

Representation for Age and Gender Group Classification	Representations for Age Estimation in Each Group		Age Estimation Error (MAE)
BIF+LSDA	Young Female	BIF+MFA	**2.66**
	Adult Female	BIF+LSDA	
	Senior Female	BIF+MFA	
	Young Male	BIF+MFA	
	Adult Male	BIF+OLPP	
	Senior Male	BIF+LSDA	

The comparisons are given in Table 7. The result from [33] is 4.66 years, which is an average of the MAEs for females (4.94 years) and males (4.38 years). In our study, various frameworks were used for completely automatic age estimation. The smallest MAE among the four representations is chosen as the representative for each framework. The error reduction rates are significant, from 36.7% to 42.9%.

Table 7. MAEs (in years) comparison of our frameworks with the result in [33] that uses manual separation of gender. Error reduction rates are computed by comparing with RPK [33].

Method	MAE	Error Reduction Rate
RPK [33]	4.66	0
F1	2.95	36.7%
F2	2.84	39.1%
F3	2.87	38.4%
F2.F	**2.66**	**42.9%**

6. Conclusions

We have systematically studied the influence of gender on age estimation using five different face representations and a large database. We also studied age estimation when performed on separate gender and age groups. Our extensive experiments show that age estimation is affected significantly by gender, and significant error reductions can be achieved if age estimation is performed on groups that each contains a small range of ages with a single gender. Based on these studies, we designed three frameworks for automatic age estimation. A data fusion approach using framework 2 gave the smallest error. It reduced the MAE by more than 40% when compared with previous results.

References

[1] *Aging of the face*. http://www.face-and-emotion.com/dataface/facets/aging.jsp.

[2] *Electronic Customer Relationship Management (ECRM)*. http://en.wikipedia.org/wiki/ECRM.

[3] D. Cai, X. He, J. Han, and H. Zhang. Orthogonal laplacianfaces for face recognition. *IEEE Trans. on Image Processing*, 15:3608–3614, 2006.

[4] D. Cai, X. He, K. Zhou, J. Han, and H. Bao. Locality sensitive discriminant analysis. In *Proc. Int. Joint Conf. on Artificial Intell.*, 2007.

[5] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. In *European Conf. on Computer Vision*, pages 484–498, 1998.

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

[7] Y. Fu, Y. Xu, and T. S. Huang. Estimating human ages by manifold analysis of face pictures and regression on aging features. In *IEEE Conf. on Multimedia and Expo*, pages 1383–1386, 2007.

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

[9] X. Geng, Z.-H. Zhou, Y. Zhang, G. Li, and H. Dai. Learning from facial aging patterns for automatic age estimation. In *ACM Conf. on Multimedia*, pages 307–316, 2006.

[10] G. Guo, Y. Fu, C. Dyer, and T. Huang. Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Trans. Image Proc.*, 17(7):1178–1188, 2008.

[11] G. Guo, Y. Fu, T. S. Huang, and C. Dyer. Locally adjusted robust regression for human age estimation. In *IEEE Workshop on Applications of Computer Vision*, 2008.

[12] G. Guo, Y. Fu, T. S. Huang, and C. Dyer. A probabilistic fusion approach to human age prediction. In *IEEE CVPR-SLAM workshop*, 2008.

[13] G. Guo, G. Mu, Y. Fu, and T. S. Huang. Human age estimation using bio-inspired features. In *IEEE CVPR*, 2009.

[14] J. Hayashi, M. Yasumoto, H. Ito, and H. Koshimizu. A method for estimating and modeling age and gender using facial image processing. In *Seventh Int. Conf. on Virtual Systems and Multimedia*, pages 439–448, 2001.

[15] Y. Kwon and N. Lobo. Age classification from facial images. *Computer Vision and Image Understanding*, 74(1):1–21, 1999.

[16] A. Lanitis, C. Draganova, and C. Christodoulou. Comparing different classifiers for automatic age estimation. *IEEE Trans. on SMC-B*, 34(1):621–628, 2004.

[17] E. Meyers and L. Wolf. Using biologically inspired features for face processing. *Int. J. Comput. Vis.*, 76:93–104, 2008.

[18] J. Mutch and D. Lowe. Object class recognition and localization using sparse features with limited receptive fields. In *Conf. on Comput. Vision and Pattern Recognit.*, pages 11–18, 2006.

[19] E. Patterson, A. Sethuram, M. Albert, K. Ricanek, and M. King. Aspects of age variation in facial morphology affecting biometrics. In *IEEE Conf. on Biometrics: Theory, Applications, and Systems*, 2007.

[20] N. Ramanathan and R. Chellappa. Face verification across age progression. *IEEE Trans. on Image Processing*, 15(11):3349–3361, 2006.

[21] K. Ricanek and E. Boone. The effect of normal adult aging on standard pca face recognition accuracy rates. In *International Joint Conf. on Neural Networks*, pages 2018–2023, 2005.

[22] M. Riesenhuber and T. Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–1025, 1999.

[23] T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio. Robust object recognition with cortex-like mechanisms. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(3):411–426, 2007.

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

[25] J. Suo, T. Wu, S. Zhu, S. Shan, X. Chen, and W. Gao. Design sparse features for age estimation using hierarchical face model. In *IEEE Int'l Conf. on Automatic Face and Gesture Recognition*, 2008.

[26] K. Ueki, T. Hayashida, and T. Kobayashi. Subspace-based age-group classification using facial images under various lighting conditions. In *IEEE conf. on FGR*, 2006.

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

[28] A. R. Webb. *Statistical Pattern Recognition, 2nd Edition*. John Wiley, 2002.

[29] S. Yan, M. Liu, and T. Huang. Extracting age information from local spatially flexible patches. In *IEEE conf. on ICASSP*, pages 737–740, 2008.

[30] S. Yan, H. Wang, T. S. Huang, and X. Tang. Ranking with uncertain labels. In *IEEE conf. on Multimedia and Expo*, pages 96–99, 2007.

[31] S. Yan, H. Wang, X. Tang, and T. Huang. Learning auto-structured regressor from uncertain nonnegative labels. In *IEEE conf. on ICCV*, 2007.

[32] 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.

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

[34] Z. Yang and H. Ai. Demographic classification with local binary patterns. In *Int. Conf. on Biometrics*, pages 464–473, 2007.

[35] S. Zhou, B. Georgescu, X. Zhou, and D. Comaniciu. Image based regression using boosting method. In *IEEE conf. on ICCV*, pages 541–548, 2005.