# Kinship Verification in the Wild: The First Kinship Verification Competition

Anonymous IJCB 2014 submission

## Abstract

*Kinship verification from facial images in wild conditions is a relatively new and challenging problem in face analysis. Several datasets and algorithms have been proposed in recent years. However, most existing datasets are of small sizes and one standard evaluation protocol is still lack so that it is difficult to compare the performance of different kinship verification methods. In this paper, we present the Kinship Verification in the Wild Competition: the first kinship verification competition which is held in conjunction with the International Joint Conference on Biometrics 2014, Clearwater, Florida, USA. The key goal of this competition is to compare the performance of different methods on a new-collected dataset with the same evaluation protocol and develop the first standardized benchmark for kinship verification in the wild.*

## 1. Introduction

Kinship verification via face images is a relatively new problem in biometrics. Compared to most existing conventional facial image analysis such as face recognition [37, 33, 34, 4, 15, 27, 20, 32, 5, 26], facial expression recognition [6, 12, 38], facial age estimation [16, 13, 17, 24, 14, 24], gender classification [28, 29] and ethnicity recognition [30, 19], there are very limited attempts on kinship verification from facial images in the literature. There are many potential applications for kinship verification such as family album organization, social media mining, and missing child search.

Recently, the performance of kinship verification by humans has been studied in psychology [2, 7, 8, 9, 21, 22], and one important observation was found: human faces can convey some important cues to identify the kin relations of persons. Inspired by this observation, computer vision researchers started to investigate the problem of kinship verification from facial images in recent years, where the objectives is to develop computational models and algorithms to verify human kin relations.

Several benchmark datasets for kinship verification are available [11, 35, 25]. However, the sizes of most existing kinship datasets are small. Moreover, one standard evaluation protocol is still lack so that it is difficult to compare the performance of different kinship verification methods. To this end, we organize the Kinship Verification in the Wild (KVW'14) Competition: the first kinship verification competition which is held in conjunction with the International Joint Conference on Biometrics 2014, Clearwater, Florida, USA. The key goal of this competition is to compare the performance of different methods on a new-collected dataset with the same evaluation protocol and develop the first standardized benchmark for kinship verification in the wild.

The remaining of this paper is organized as follows: Section 2 overviews the existing works on kinship verification via face images. Section 3 introduces the newly collected dataset and experimental protocol. The baseline method and results are presented in Section 4. Section 5 presents the evaluation results of all participants' methods. Section 6 summarizes the results obtained by different participants of the competition. Finally, Section 7 concludes the paper.

## 2. Overview of Existing Works

Over the past five years, several kinship verification via face images approaches have been proposed in computer vision and biometrics [11, 35, 31, 39, 36, 18, 40, 23, 10, 25]. Generally, these methods can be categorized into two streams: descriptor-based [11, 39, 40, 18] and similarity learning-based [35, 36, 31, 25]. For descriptor-based methods, some important cues such as skin color [11], histogram of gradient [11], Gabor gradient orientation pyramid [40], salient part [23], self-similarity [18], and dynamic expressions [10], are usually employed for face representation. For similarity learning-based methods, subspace and metric learning are used to learn a semantic feature space to better measure the similarities of face samples. Representative such algorithms include transfer subspace learning [36] and neighborhood repulsed metric learning [25]. Table 1 reviews existing kinship verification methods which were presented over the recent five years, where their performance is evaluated by the mean verification rate. While the performance of different methods cannot be compared directly because of different datasets and protocols, we still see that there has been substantial improvement in kinship verifica-

Table 1. Performance comparison of recent kinship verification methods.

| Method | Feature representation | Classifier | Dataset | Accuracy (%) | Year |
|---|---|---|---|---|---|
| Fang *et al.* [11] | Local features of face parts | KNN | Cornell KinFace | 70.7 | 2010 |
| Xia *et al.* [35] | Transfer subspace learning | KNN | UB KinFace | 60.0 | 2011 |
| Zhou *et al.* [39] | Spatial pyramid local feature | SVM | 400 pairs (N.A.) | 67.8 | 2011 |
| Xia *et al.* [36] | Context feature with transfer learning | KNN | 296 pairs (N.A.) | 79.9 | 2012 |
| Kohli *et al.* [23] | Self similarity of Weber face | SVM | 272 pairs (N.A.) | 74.1 | 2012 |
| Lu *et al.* [25] | Local feature with metric learning | SVM | KinFaceW-I / II | 69.9 / 76.5 | 2012 |
| Dibeklioglu *et al.* [10] | Dynamic spatio-temporal appearances | SVM | 228 pairs (N.A.) | 72.9 | 2013 |

tion in recent years. Moreover, we believe there is considerable space for further improvement.

## 3. Dataset and Protocol

In this competition, we collected a large face kinship dataset by the online web search, where several hundreds of public figures' face images and their parents' or children's face images were crawled. The face images were collected without restriction in terms of pose, expression, illumination, background, age, ethnicity, and occlusion. We define kinship as a relationship between two persons who are biologically related with overlapping genes. Therefore, we examine four different kin relations: Father-Son (F-S), Father-Daughter (F-D), Mother-Son (M-S) and Mother-Daughter (M-D). We provided three sets in this competition: training set, validation set and testing set. In the training set, we construct 300 positive and 300 negative pairs of kinship images for each of the four relations. In the validation set, we provided 50 positive and 50 negative pairs of kinship images for each of the four relations. In the testing phase, we provided 600 image pairs for each kinship relation for evaluation. There is no overlap between different sets. For ease of use, we manually labeled the coordinates of the eyes position of each face image, and cropped and aligned facial region into $64 \times 64$ so that the competition participants can focus more on the kinship verification algorithms development rather than face alignment because face images in our dataset were captured in the wild and it is challenging to precisely localize facial fiducial points. Figures 1 and 2 show some positive and negative image pairs for different kin relation in our dataset, where images from top to bottom are from the F-S, F-D, M-S and M-D kin relations, respectively.

Generally, there are two protocols in verification tasks: closed-set and open-set [3]. In this competition, we designed an open-set verification protocol because we expect the designed kinship verification systems can verify where there is a kinship relation for a new face pair without redesigning the verification system. Specifically, the training set is used to learn the model and the validation set is employed to tune the parameters of the models. The testing set is used to evaluate the generalization capability of the developed kinship verification methods. The verification rates and receiver operating characteristic (ROC) curves of different kinship verification methods are compared for evaluation.

## 4. Baseline Results

In our competition, we provide a baseline method which uses the LBP feature representation and the cosine similarity for kinship verification. For each face image, we densely sampled $16 \times 16$ blocks with the stepsize of 8 pixels, and we can obtain 49 blocks in each whole face. For each block, we extracted a 59-dimensional uniform pattern histogram feature by following [1] to describe each image block. Then, we concatenated features extracted in all blocks to form a 2891-dimensional feature vector for final feature representation. Figure 3 shows the verification rate and ROC curve of our baseline method.

## 5. Participants' Results

In total, four participants contributed to the competition. Then, we briefly describe the submitted methods.

Kou *et al.* proposed a similarity learning based kinship verification method. They used the HOG feature descriptor to describe each face image. Specifically, each face image was divided into $8 \times 8$ non-overlapped blocks and the size of each block is $8 \times 8$. For each block, they extracted a 9-dimensional histogram feature. Then, they concatenated the features extracted from each block into a 576-dimensional feature vector for face representation. In order to effectively measure the kin similarity for a given pair of facial images, they proposed to explicitly learn a similarity function instead of the commonly-used distance metric. Specifically, the similarity function was represented by a bi-linear function parameterized by a transform matrix $\mathbf{W}$, which is not necessary to be semi-positive define or symmetric. The objective to learn $\mathbf{W}$ is to minimize a hinge loss of the labeled triplets from the training set, combined with a low-rank regularization of $\mathbf{W}$. Finally, $\mathbf{W}$ is obtained by a stochastic gradient descent algorithm.

Castrillón-Santana *et al.* proposed a local feature based kinship verification approach. Specifically, they used three

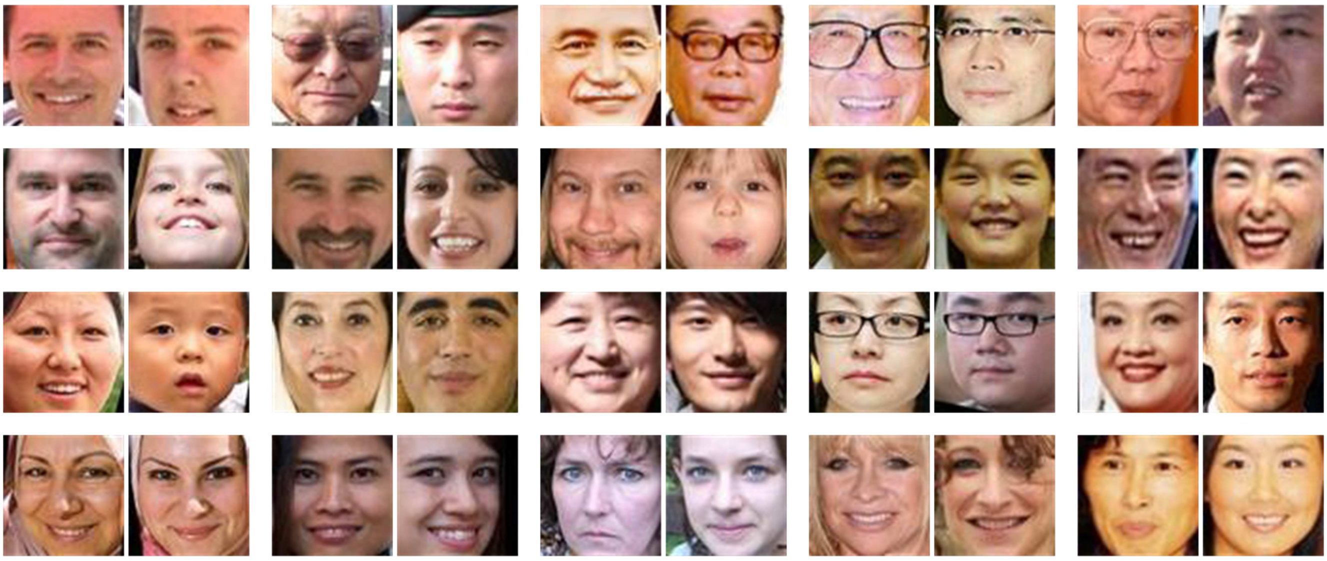


Figure 1. Several positive examples of our dataset. From top to bottom are the F-S, F-D, M-S and M-D kin relations, and the neighboring two images in each row are with the kin relation, respectively.

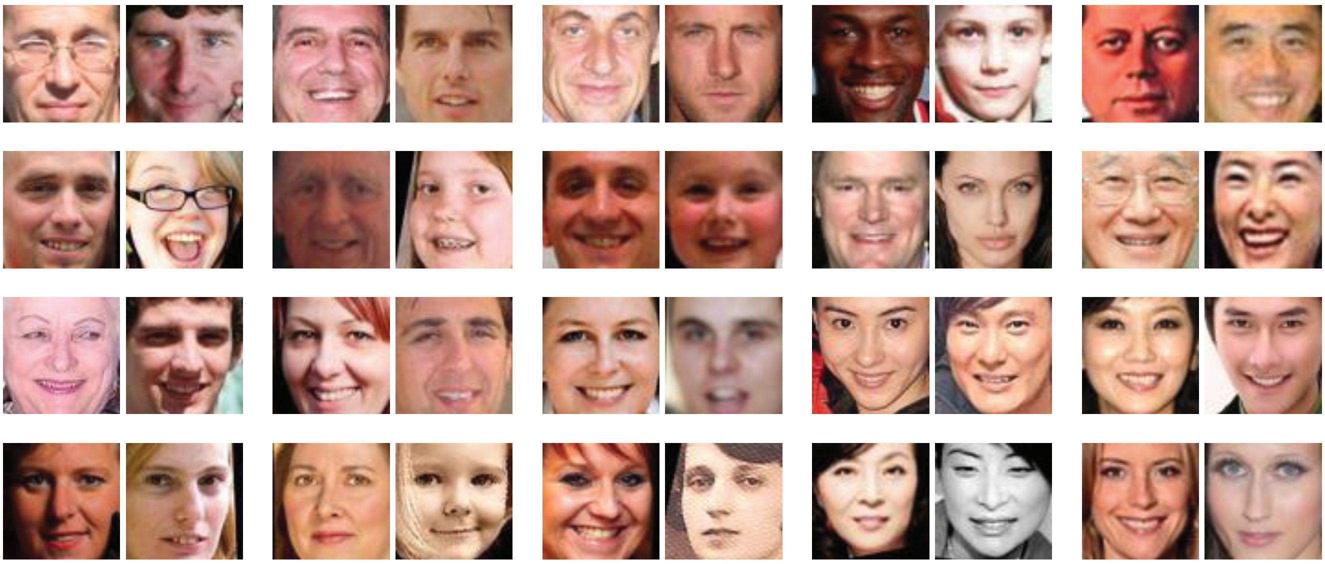
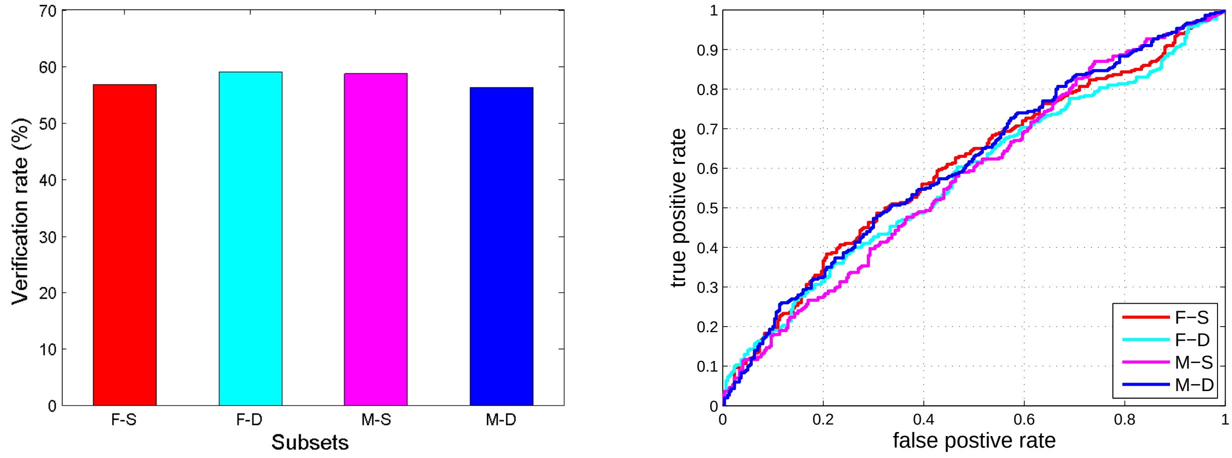


Figure 2. Several negative examples of our dataset. From top to bottom are the F-S, F-D, M-S and M-D kin relations, and the neighboring two images in each row are without the kin relation, respectively.

local features including LBP, Local Salient Patterns and HOG for face representation, where each face image was divided into $5 \times 5$, $5 \times 5$, and $8 \times 8$ blocks, respectively. The intersectional kernel was employed to computer the similarity of each face pair. Finally, the SVM classifier was used for classification.

Bottino *et al.* proposed an attribute combination method for kinship verification, where geometric and both global and local textural features are defined as the attributes. In their method, Planar Projection Summation Invariants (PPSI), Weber Local Descriptor (WLD) and SIFT features extracted for each face image. They further employed three feature selection methods including the minimum-Redundancy-Maximum-Relevance (mRMR), optimal mRMR, and the modified Sequential Forward Floating Selection method to select the most informative features. Finally, the SVM classifier was used for classification.

Ghahramani *et al.* proposed a local feature based kinship verification approach. They employed Uniformly-sampled Thresholds for LBP (UTLBP) to extract features from faces. Since conventional LBP does not fully capture the detailed information of the relative pixel information, UTLBP can extract information of the surrounding pixels intensity to the centre pixel to reflect facial similarity among faces in a fam-




(a) Verification rate (b) ROC curve

Figure 3. The verification rate and ROC curve of our baseline method, respectively.

ily. They used different thresholds in the step function. The vector size of each LBP was also reduced by implementing LMNN and selection of the top 20 features. Another shortcoming of using histograms is lack of spatial location. Due to low resolution of photos, they divided the image into four rectangles by using the Perpendicular bisector of the $x$ and $y$ dimension. The top 20 features of UTLBP descriptor are then concatenated from each of four divisions. The threshold adjustment is the uniform sampling in the range $[-50, 50]$ with the step of 25. Hence they got 200 features in the end for each face that is smaller than using the conventional LBP on the whole face. The scores were calculated using SVM.

Table 2 tabulates the verification rates of different participants on our kinship dataset, and Figure 4 shows the ROC curves of different participants obtained on different subsets, where "CNU", "ULPGC", "POT", and "Oulu" denote Capital Normal University, Universidad de Las Palmas de Gran Canaria, Politecnico di Torino, and University of Oulu, respectively. According to the results shown in Table 2 and Figure 4, we are pleased to announce that the winners of this competition are the participants from CNU and ULPGC as they achieved the same mean verification rate. Moreover, one of them achieved the best verification rate on two subsets (F-S and M-D) and another obtained the best results on the other two subsets (F-D and M-S).

## 6. Discussion

The first kinship verification in the wild competition has been a great community effort. We expect to have established a new benchmark for kinship verification via face images, which will allow researchers in this field to further investigate this problem. To keep this benchmark available in the future, the KVW 2014 organizers are keeping the dataset available through their online repository, and they will continue to update the new progress on this dataset in the future.

One important message to convey in this competition is that one can learn what are the current trends and state-of-the-arts in this field. For instance, three teams participated in this competition used the SVM classifier and the other one used the similarity learning technique. According to the results, it is derisible to combine both of these techniques to further improve the verification performance.

## 7. Conclusion

This paper describes the kinship verification in the wild competition: the first kinship verification competition held in conjunction with the International Joint Conference on Biometrics 2014, Clearwater, Florida, USA. The main challenge of the competition is to verify whether there is a kin relation for a given pair of face images which were captured in the wild. In this competition, the largest face kinship dataset is provided and a standard protocol and benchmark is presented. In total four participants submitted to this competition, we can see that current technology is still not enough to produce reasonably good results and there is much space for further improvement.

## References

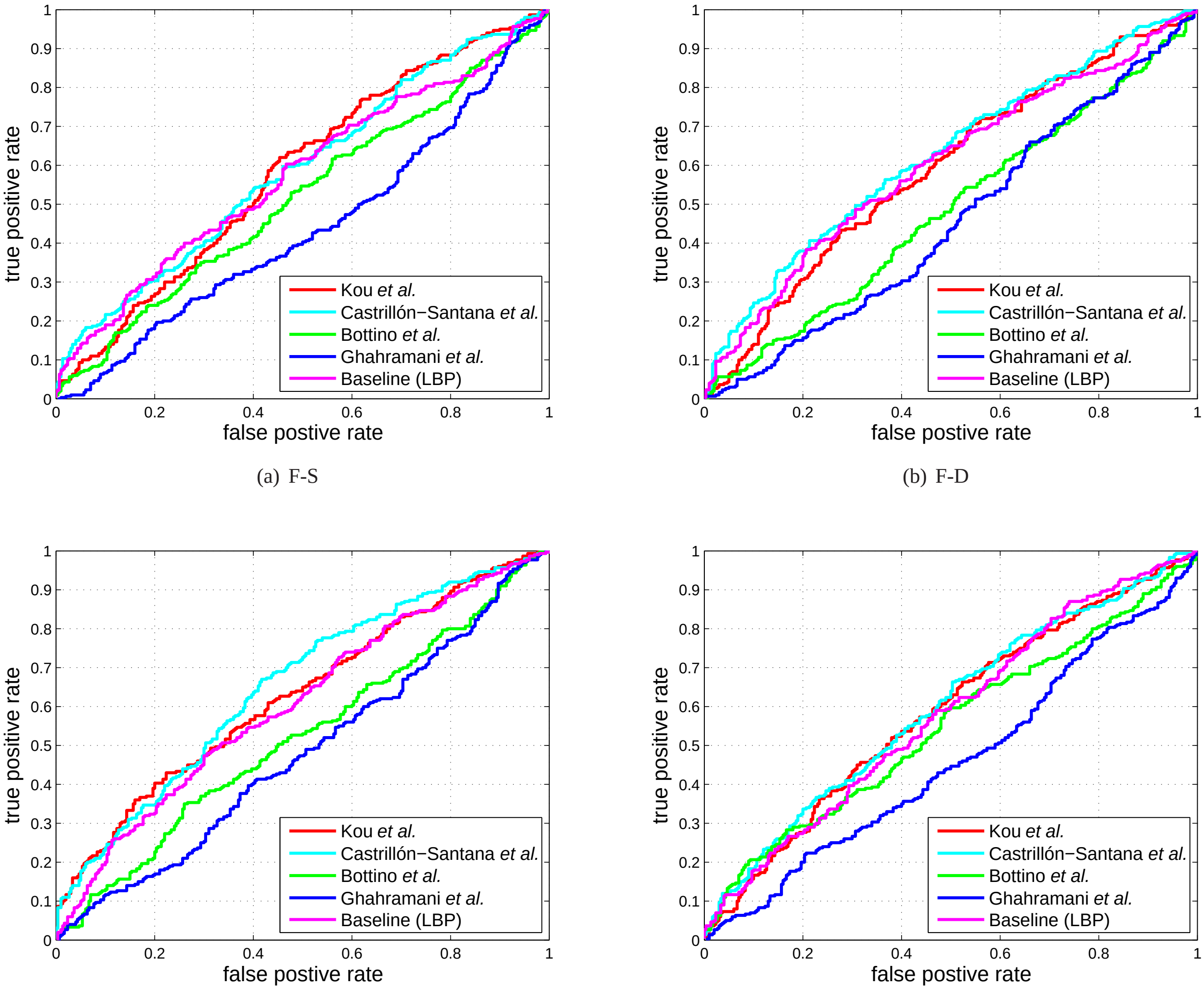
[1] T. Ahonen, A. Hadid, et al. Face description with local binary patterns: application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006.

[2] A. Alvergne, R. Oda, C. Faurie, A. Matsumoto-Oda, V. Durand, and M. Raymond. Cross-cultural perceptions of facial resemblance between kin. *Journal of Vision*, 9(6), 2009.

IJCB 2014

Table 2. The verification rates (%) of different participants on our kinship dataset.

| Rank | Authors | Affiliation | F-S | F-D | M-S | M-D | Mean |
|---|---|---|---|---|---|---|---|
| 1 | Kou *et al.* | CNU | **58.2** | 58.0 | 60.3 | **57.0** | **58.4** |
| 1 | Castrillón-Santana *et al.* | ULPGC | 56.0 | **59.8** | **61.7** | 55.8 | **58.4** |
| 3 | Bottino *et al.* | POT | 54.7 | 56.8 | 58.8 | 55.2 | 56.4 |
| 4 | Ghahramani *et al.* | Oulu | 50.5 | 50.2 | 51.3 | 50.8 | 50.7 |
| - | Baseline | - | 56.8 | 59.0 | 58.7 | 56.3 | 57.7 |




(a) F-S (b) F-D (c) M-S (d) M-D

Figure 4. The ROC curves of different participants obtained on the (a) F-S, (b) F-D, (c) M-S, and (d) M-D subsets, respectively.

[3] E. Bailly-Bailliére, S. Bengio, F. Bimbot, M. Hamouz, J. Kittler, J. Mariéthoz, J. Matas, K. Messer, V. Popovici, F. Porée, et al. The banca database and evaluation protocol. In *International Conference on Audio-and Video-Based Biometric Person Authentication*, pages 1057–1057, 2003.

[4] P. N. Belhumeur, J. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[5] Z. Cao, Q. Yin, X. Tang, and J. Sun. Face recognition with learning-based descriptor. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 2707–2714, 2010.

[6] I. Cohen, N. Sebe, A. Garg, L. Chen, and T. Huang. Facial expression recognition from video sequences: Temporal and static modeling. *Computer Vision and Image Understanding*, 91(1):160–187, 2003.

[7] M. Dal Martello and L. Maloney. Where are kin recognition

signals in the human face? *Journal of Vision*, 6(12), 2006.

[8] M. Dal Martello and L. Maloney. Lateralization of kin recognition signals in the human face. *Journal of Vision*, 10(8), 2010.

[9] L. DeBruine, F. Smith, B. Jones, S. Craig Roberts, M. Petrie, and T. Spector. Kin recognition signals in adult faces. *Vision Research*, 49(1):38–43, 2009.

[10] H. Dibeklioglu, A. A. Salah, and T. Gevers. Like father, like son: Facial expression dynamics for kinship verification. In *IEEE International Conference on Computer Vision*, pages 1497–1504, 2013.

[11] R. Fang, K. Tang, N. Snavely, and T. Chen. Towards computational models of kinship verification. In *IEEE International Conference on Image Processing*, pages 1577–1580, 2010.

[12] B. Fasel and J. Luettin. Automatic facial expression analysis: a survey. *Pattern Recognition*, 36(1):259–275, 2003.

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

[14] X. Geng, Z. Zhou, and K. Smith-Miles. Automatic age estimation based on facial aging patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2234–2240, 2007.

[15] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.

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

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

[18] G. Guo and X. Wang. Kinship measurement on salient facial features. *IEEE Transactions on Instrumentation and Measurement*, 61(8):2322–2325, 2012.

[19] S. Gutta, J. Huang, P. Jonathon, and H. Wechsler. Mixture of experts for classification of gender, ethnic origin, and pose of human faces. *IEEE Transactions on Neural Networks*, 11(4):948–960, 2002.

[20] X. He, S. Yan, Y. Hu, P. Niyogi, and H. J. Zhang. Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(3):328–340, 2005.

[21] G. Kaminski, S. Dridi, C. Graff, and E. Gentaz. Human ability to detect kinship in strangers' faces: effects of the degree of relatedness. *Proceedings of the Royal Society B: Biological Sciences*, 276(1670):3193–3200, 2009.

[22] G. Kaminski, F. Ravary, C. Graff, and E. Gentaz. Firstborns' disadvantage in kinship detection. *Psychological Science*, 21(12):1746–1750, 2010.

[23] N. Kohli, R. Singh, M. Vatsa, and I. IIIT-Delhi. Self-similarity representation of weber faces for kinship classification. In *International Conference on Biometrics Theory, Applications and Systems*, 2012.

[24] A. Lanitis. Evaluating the performance of face-aging algorithms. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 1–6, 2008.

[25] J. Lu, J. Hu, X. Zhou, Y. Shang, Y.-P. Tan, and G. Wang. Neighborhood repulsed metric learning for kinship verification. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 2594–2601, 2012.

[26] J. Lu, Y.-P. Tan, and G. Wang. Discriminative multimanifold analysis for face recognition from a single training sample per person. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):39–51, 2013.

[27] A. Martínez and A. Kak. PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):228–233, 2002.

[28] B. Moghaddam and M. Yang. Gender classification with support vector machines. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 306–311, 2000.

[29] B. Moghaddam and M. Yang. Learning gender with support faces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):707–711, 2002.

[30] Y. Ou, X. Wu, H. Qian, and Y. Xu. A real time race classification system. In *IEEE International Conference on Information Acquisition*, pages 1–6, 2005.

[31] M. Shao, S. Xia, and Y. Fu. Genealogical face recognition based on ub kinface database. In *Intenrational Conference Computer Vision and Pattern Recognition Workshops*, pages 60–65, 2011.

[32] X. Tan, S. Chen, Z. Zhou, and F. Zhang. Face recognition from a single image per person: A survey. *Pattern Recognition*, 39(9):1725–1745, 2006.

[33] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[34] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):775–779, 1997.

[35] S. Xia, M. Shao, and Y. Fu. Kinship verification through transfer learning. In *International Joint Conference on Artifical Intelligence*, pages 2539–2544, 2011.

[36] S. Xia, M. Shao, J. Luo, and Y. Fu. Understanding kin relationships in a photo. *IEEE Transactions on Multimedia*, 14(4):1046–1056, 2012.

[37] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: a literature survey. *ACM Computing Surveys*, 35(4):399–458, 2003.

[38] W. Zheng, X. Zhou, C. Zou, and L. Zhao. Facial expression recognition using kernel canonical correlation analysis (kcca). *IEEE Transactions on Neural Networks*, 17(1):233–238, 2006.

[39] X. Zhou, J. Hu, Y. Lu, J. Shang, and Y. Guan. Kinship verification from facial images under uncontrolled conditions. In *ACM International Conference on Multimedia*, pages 953–956, 2011.

[40] X. Zhou, J. Lu, J. Hu, and Y. Shang. Gabor-based gabor-based gradient orientation pyramid for kinship verification under uncontrolled environments. In *ACM International Conference on Multimedia*, pages 2322–2325, 2012.