Understanding Kin Relationships in a Photo

Siyu Xia, Ming Shao, *Student Member, IEEE*, Jiebo Luo, *Fellow, IEEE*, and Yun Fu, *Senior Member, IEEE*

***Abstract*—There is an urgent need to organize and manage images of people automatically due to the recent explosion of such data on the Web in general and in social media in particular. Beyond face detection and face recognition, which have been extensively studied over the past decade, perhaps the most interesting aspect related to human-centered images is the relationship of people in the image. In this work, we focus on a novel solution to the latter problem, in particular the kin relationships. To this end, we constructed two databases: the first one named UB KinFace Ver2.0, which consists of images of children, their young parents and old parents, and the second one named FamilyFace. Next, we develop a transfer subspace learning based algorithm in order to reduce the significant differences in the appearance distributions between children and old parents facial images. Moreover, by exploring the semantic relevance of the associated metadata, we propose an algorithm to predict the most likely kin relationships embedded in an image. In addition, human subjects are used in a baseline study on both databases. Experimental results have shown that the proposed algorithms can effectively annotate the kin relationships among people in an image and semantic context can further improve the accuracy.**

***Index Terms*—Context, face recognition, kinship verification, semantics.**

I. Introduction

WITH the development of technology in modern multimedia society, image acquisition and storage by digital devices have never been easier than today. Storage unit like GB or TB is not qualified already in storing images from the Internet. For example, as the most popular social network website around the world, Facebook has already hosted over 20 billion images, with more than 2.5 billion new photos being added each month [1]. However, how to successfully and automatically manage the substantial images captured by people is a real challenge since it pushes the computer to its limit of image understanding—it requires both large-scale data analysis and high accuracy. In most cases, people are the focus of images taken by consumers and managing or organizing them essentially raises two problems: 1) who these people are and 2) what their relationships are.

Manuscript received June 26, 2011; revised December 05, 2011; accepted January 18, 2012. Date of publication February 10, 2012; date of current version nulldate. This work was supported in part by the IC Postdoctoral Research Fellowship under Award 2011-11071400006, AFOSR Award FA9550-12-1-0201, and by Google Faculty Research Awards. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Qi Tian.

S. Xia is with the School of Automation, Southeast University, Nanjing 210096, China (e-mail:xia081@gmail.com).

J. Luo is with the Computer Science Department, University of Rochester, Rochester, NY 14627 USA (e-mail: jiebo.luo@rochester.edu).

M. Shao and Y. Fu are with the Department of Computer Science and Engineering, State University of New York at Buffalo, Buffalo, NY 14260 USA (e-mail: mingshao@buffalo.edu; yunfu@buffalo.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2012.2187436

In the first problem, identities are what we are most concerned with and intuitively faces are critical clues. Face recognition is therefore an important approach toward this problem. Basically, face recognition can be further categorized into two classes according to whether the contextual information is used. Face recognition without context information, namely, pairwise face recognition, has been extensively studied during the past decade by exploration of following techniques: face detection and alignment [2], [3], subspace learning [4]–[10], invariant feature extraction [11]–[13], metric learning method [14], attributes based classifier [15] and face synthesis and hallucination [16], [17]. Nevertheless, it is still an ongoing problem due to several practical factors, e.g., illumination, pose, expression and aging. The performance of the face recognition algorithm is dramatically degraded when a large-scale database is considered [18]. On the other hand, in practice, faces no longer appear alone due to the rapid development of multimedia. What often accompany faces are text, video and many other metadata. Recently, research attention is shifting to contextual information involved in the people-centric images, including locations, capture time of images and patterns of cooccurrence and reoccurrence in images [19], social norm and conventional positioning observed [20], [21], text or other linked information [22], [23], and clothing [24].

The relationship of people in a photo also deserves the research attention in that social connection has been the essence of modern society. The possible relationships in consumer images include "kinship," "friends," "colleague," etc. Statistically, such relationships are often the main themes of people's albums and successfully parsing or tagging them leads to better understanding of images. Among them, kinship is believed to be the most discriminative one since children naturally inherit gene from their parents. An intuitive way to annotate this relationship is by identities of individuals the images contain, and in theory, this can be achieved using automatic face recognition—the first problem aforementioned. However, it is also possible that the relationship is still uncertain even if identities are known. Therefore, direct kinship verification is worth investigating. The pioneer work in [25] and [26] attempted to discriminate kinship based on selected inheritance—invariant features. When kin relationship, gender and age are known, a family tree can be automatically created given an family image like Fig. 1. Moreover, kinship verification can be applied to both general computer vision problems, such as image retrieval or annotation, and some specific application scenarios, such as finding missing children.

In this paper, as an extension of our previous work [27], [28], we attempt to tackle the kinship discrimination problem based on faces as well as the semantics embedded in contexts. This is reasonable because both appearance and semantics are valuable for relationship inference. Recent research [29]–[31] discovers that facial appearance is a cue for genetic similarity as children resemble their parents more than other adults of the

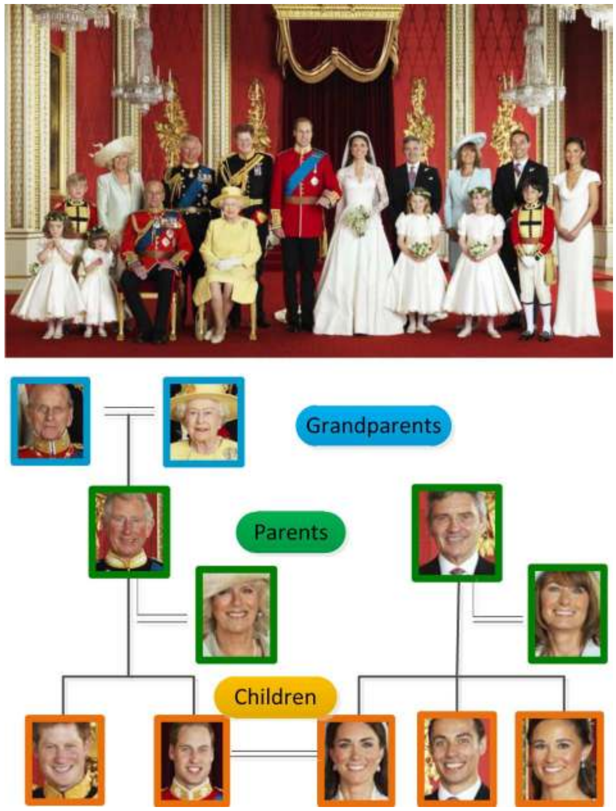


Fig. 1. A family tree (bottom) can be extracted from a family photo (top) if kin relationships among people in the photo are known. The double horizontal line shown in the figure represents the couple relationship.

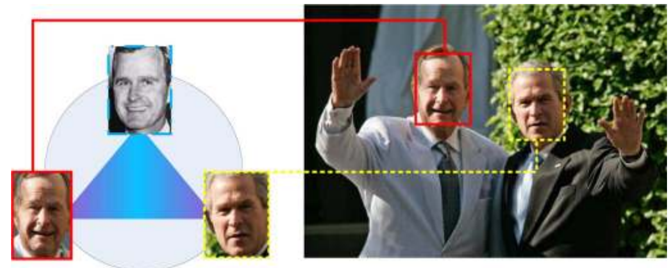


Fig. 2. Faces of parents captured while they were young (top of the blue triangle in online version) are more alike their children's (lower right corner of the blue triangle in online version) compared with images captured when they are old (lower left corner of the blue triangle in online version). By leveraging the intermediate set, parent–child verification problem becomes more tractable.

same gender, and that there is differential resemblance between two parents, depending on the age and gender of the child. Analogously, a critical observation is that faces of parents captured while they were young closely resemble their children's compared with images captured when they are old, as Fig. 2 illustrated. This promising statistics inspires us to perform the following research. First, the UB KinFace[1] Ver2.0 database [28] is set up by collecting child, young–parent and old–parent face images from the Web. Through this database, aforementioned hypothesis based on genetics theory is preliminarily proved. Second, a transfer subspace learning based method is proposed, aiming at reducing the large divergences of distributions between children and old parents. The key idea is to utilize some intermediate data set close to both the source and target distributions and naturally the young–parent set is suitable for this task. Third, to exploit the value of contextual information and semantics in kinship, a database called FamilyFace is built and its images are drawn from social network websites, such as Facebook and Flickr. Then we propose a strategy to infer the most possible kin relationships embedded in the images. In addition, to conduct both objective and subjective evaluations, a human based test is introduced as a baseline and a more systematic comparison is performed as well. The difference between this paper and our previous works [25], [26] lie in the following: 1) we utilize young parents set as an intermediate set to reduce the divergence between data distributions and 2) we consider the context and semantics in family photos and treat the kinship verification as a joint tagging problem, mining more knowledge than mere pairwise verification.

[1]UB KinFace Ver2.0 (for noncommercial purpose) now is available at: http://www.cse.buffalo.edu/~yunfu/research/Kinface/Kinface.htm

The rest of the paper is organized as follows. Several related works are reviewed in Section II. In Section III, we present the transfer subspace learning method for kinship verification. Both contextual and semantic information are utilized to further enhance the accuracy. Experimental evaluations are described in Section IV. Finally, we draw conclusions in Section V.

II. Related Work

The first attempt of kinship verification was published in [25] and [26]. To utilize local information, they first localized key parts of faces, so that facial features, such as skin color, gray value, histogram of gradient, and facial structure information, are extracted. Then K-nearest-neighbor (KNN) with Euclidean metric was adopted to classify faces. Such simple strategy works reasonably well for online collected data. Different from the verification problem of the same person, kinship verification may not expect each feature pair is exactly the same, because genetic variations often exist from parents to children. Moreover, similarities or features on faces between kins are mostly located at eyes, nose, mouth, etc., according to genetics studies. We therefore extract these local features inherited from parents rather than holistic ones. Though considering this intuitive information, the kinship verification is still challenging due to remarkable differences between the query and gallery images. The most significant degrading factor, in terms of face recognition, is *aging*. Parent images used in kinship verification often contain the elders, but queries are usually young males or females. Texture distributions of these faces are quite different due to the aging effect, let alone the structure variations on faces from different identities. Meanwhile, other uncontrollable factors are evident in the real-world, as described in [18]. These factors together lead to a complex new problem in biometrics.

Moreover, recently a more significant problem draws considerable attentions that common assumption of training and test data from the same feature space and distribution is not always fully valid. This is a natural situation for any new classification problems. Manual labeling work is time-consuming and people want to reuse the knowledge that has already been extensively studied. In such a case, knowledge transfer or transfer learning is highly desirable. In [32], several practical instances have been introduced to illustrate the role of transfer learning, e.g., web-document classification, WiFi data remodeling and sentiment classification based on customer reviews. The basic problem is how to reuse the knowledge learned from other data or feature spaces. Specifically, transfer learning as we mentioned here can be further categorized into two classes, inductive transfer learning [33]–[38] and transductive transfer learning [39]–[41]. For the former one, the domains in which two sets of data embedded are either the same or different, but learning targets are always different. Meanwhile, the latter one can tolerate different distributions between data sets, but learning targets are quite identical. In this paper, for our kinship verification problem, three distributions exist, i.e., children, young parents and old parents. Different from [25] and [26], we introduce young–parent data set as an intermediate distribution to facilitate transfer learning and clearly our approach falls into the transductive transfer learning category.

Bregman divergence was adopted in [42] to minimize the distribution divergence in the subspace by seeking a common subspace. This work has been applied to the cross-domain age estimation [43] and improved the accuracy on one data set by the knowledge transferred from the other data set. Sometimes, however, transfer learning may fail due to the large discrepancy between the source and target data sets [32], such as children and old parents images. We therefore consider introducing an intermediate data set as a link to abridge the divergence between the source and target domains. These intermediate samples, i.e., face images of young parents, are close to both children and old parents, and appropriately designed for our formulation.

Understanding human characteristics in images by contextual information has been extensively investigated in [19]–[24], [44], and [45] to further assist image understanding. In [19], the time stamp, location of image, reoccurrence, cooccurrence among image collections and annotated individuals were leveraged to infer unknown people. This work highly relies on the assumption that people in image collections may emerge more than once, be tied to each other, and with high likelihood stand in the same location with the same people. Relations between people's positions in images, and their genders/ages were revealed in [20]. This can be explained by the fact that people's positions in photos indicate social settings in real-world. Moreover, these factors together, e.g., position, age, and gender, can also help human identification in a family photo [21]. In addition, [22] and [23] addressed the problem that when there are no labels or poor labels in training set, how to annotate faces by text-image cooccurrence, which is naturally contained in Web resources. It was argued in [24] that identities of people are highly related to their clothing, and clothing segmentation and face recognition are intertwined. Therefore, clothing features can be used to strengthen discriminative feature extraction. Face clustering technology has also been used in [44] to discover social relationships of subjects in personal photo collections. Cooccurrence of identified faces as well as interface distances (inferred from the in-image distance and typical human face size) were used to calculate the link strength between any two identified persons. Last but not least, Markov logic [45] was utilized to formulate the rules of thumb as soft constraints, when a group of photos are considered collectively with the purpose of detecting relationships in consumer photo collections.

III. Proposed Approach

A. UB KinFace Database

To the best of our knowledge, the "UB KinFace Ver1.0" [27] is the first database containing images of both children and their parents at different ages. All images in the database are real-world collections of public figures (celebrities and politicians) from the Web. We use a person's name as the query for image search. The database used in this paper, named "UB KinFace Ver2.0" [28], is an extension of "UB KinFace Ver1.0" by considering more instances and ethnicity group impacts. A general view of this database is summarized in Fig. 3. Similar to "Labeled Faces in the Wild" [18], these unconstrained facial images show a large range of variations, including pose, lighting, expression, age, gender, background, race, color saturation, and image quality, etc.

The key difference between our database and that in [25] lies in our inclusion of young parents. Basically, our database comprises 600 images of 400 people which can be separated into 200 groups. Each group is composed of child, young–parent and old–parent images. The "UB KinFace Ver2.0" can be mainly divided into two parts in terms of race, i.e., Asian and non-Asian, each of which consists of 100 groups, 200 people and 300 images. Typically, there are four kin relations, i.e., "son–father," "son–mother," "daughter–father," and "daughter–mother," as shown in Fig. 3. In addition, Fig. 3 illustrates the statistics from the perspective of race and kin relations. As we can see, male celebrities of both Asian and non-Asian, either son or father, are dominant in the UB KinFace database. When four possible kin relations are considered in our problem, "fathers" reasonably become the essential roles, with 46.5% and 38.5% over all groups in "son–father" and "daughter–father" relations. This phenomenon can be explained by the fact that there are statistically more notable males than females in current government, entertainment and sports communities.

Contexts and semantics in family albums are other factors that we can leverage to improve the accuracy. To this end, a new database called FamilyFace is built, which, including 214 images, 507 persons in total, is collected from the popular social network websites, e.g., Facebook and Flickr. The example images are shown in Fig. 4. In this database, there are kinship and other relationships, e.g., friendship and colleague, in each image, and they may coexist in the same image. In addition, not only Asian but also western people are considered in the database. Since we downloaded them from the Web and they are not subject to any specific constraints, these images may reveal real-world conditions for practical evaluations.

B. Feature Extraction

Two typical features that can distinguish true child–parent pairs from the false ones are explored in this section. One is based on appearance and extracted by Gabor [11] filters (eight directions and five scales). Particularly, we first partition each face into regions in five layers, as illustrated in Fig. 5. As we can see, the entire face is the first layer. The second layer includes upper, lower, left, right and center parts of the face. The forehead, eyes, nose, mouth, and cheek areas constitute the third layer and their finer sub-regions form the fourth layer. A group of sub-regions based on the four fiducial points finally form

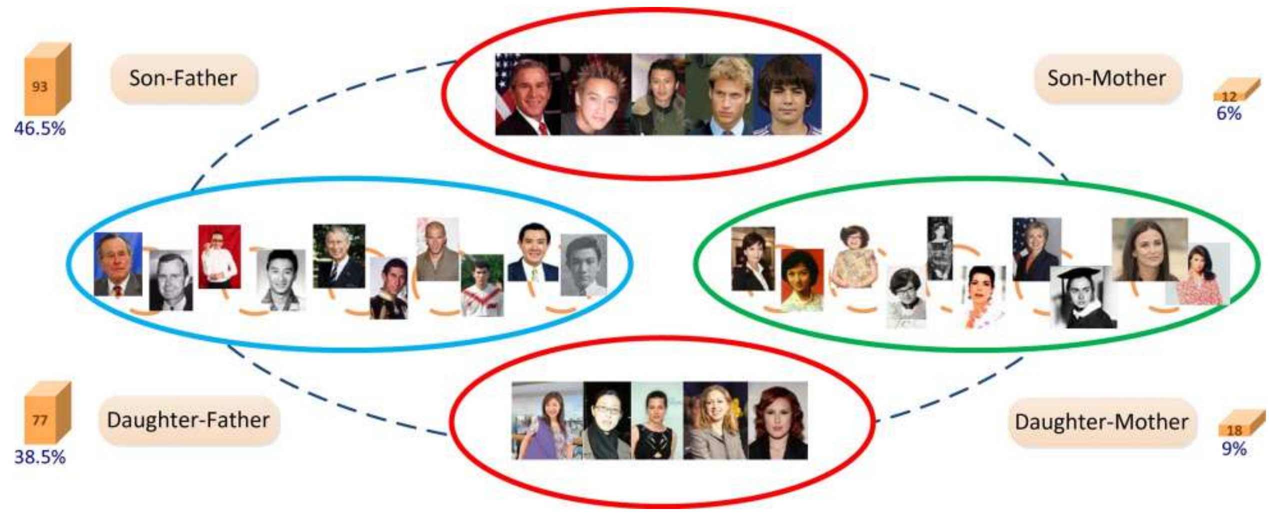


Fig. 3. KinFace database illustrations. Four groups of people and the corresponding four possible relationships: son–father, son–mother, daughter–father, and daughter–mother. Children images are in the red ellipse; male parents images are in the blue ellipse; female parents images are in the green ellipse (in online version).

Fig. 4. Sample images (top) from the FamilyFace database. The cropped faces contained in these images are shown in the second row.

the fifth layer. Then we impose Gabor filters on each local region. Similar idea has been adopted in [46] to analyze facial expressions through local features. Intuitively, kinship verification is also a process on local regions. For instance, when people are talking about kinship, they often compare regions on faces between children and their parents and wonder whether they share similar eyes, noses, or mouths. Another feature is based on the anthropometric model [47] which essentially considers structure information of faces. Based on the captured key points, we obtain 6-D structure features of ratios of typical region distances, e.g., "eye–eye" distance versus "eye–nose" distance. Structured information is believed to inherit largely from parents, and therefore might be promising for kinship verification [25]. However, due to aging process [47], the old parents face structures are deformed from the ones when they were young. So, we use transfer subspace learning to mitigate this degrading factor.

Since images obtained from the Web are under arbitrary environment, we first take advantage of "total variation" to remove lighting effects. Total variation, as first used in image denoising, has been successfully applied to illumination free face recognition [48]. After removing the lighting effect, we partition faces according to key points, i.e., the locations of left eye, right eye, nose tip, and the center of the mouth, into regions with different widths and heights, as shown in Fig. 5.

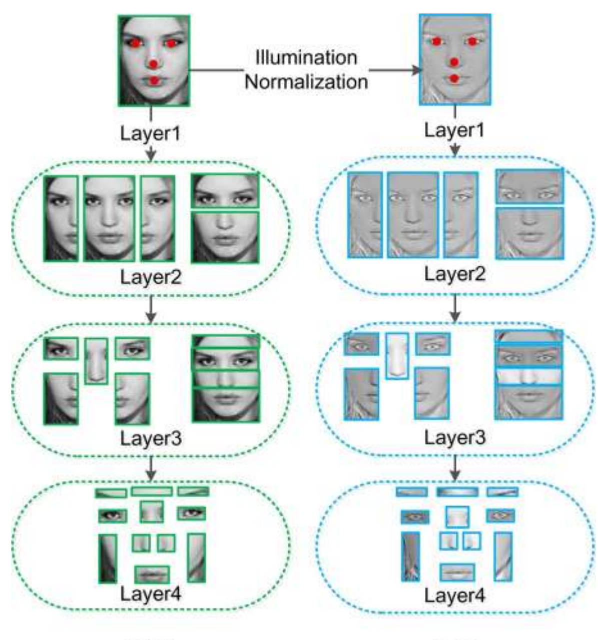


Fig. 5. Face partitions in different layers and face image illumination normalization. For simplicity, only four layers are illustrated instead of five. Red dots on faces in Layer1 illustrate four key points mentioned in this paper (in online version).

To fully utilize all the features in different regions, we propose a new feature extraction strategy based on cumulative match characteristic (CMC) [49]. We add features in several rounds and at each round we select the feature in one region that can maximize the difference of recognition performance of child–old parent between current round and last round. This process can be formulated as

$$\max\{\mathrm{cum}_{i+1}(S_{\mathrm{CvsO}}) - \mathrm{cum}_i(S_{\mathrm{CvsO}})\} \tag{1}$$

where $\mathrm{cum}_i(S_{\mathrm{CvsO}})$ means the sum of recognition rates from rank 1 to n, and C and O denote child and old–parent groups,

TABLE I
SELECTED REGIONS IN DIFFERENT LAYERS

Layers	L1	L2	L3	L4	L5
Regions	1	$1,2,3$	$1,3$	$2,3,7,8,11$	$14-30$

respectively. Intuitively, this strategy intends to choose the feature that best discriminates true child–parent pairs from false ones. In addition, to accelerate the feature selection process, we choose candidate features in the current layer until no more can be added to enhance the performance. Then we will repeat this process in the next layer and perform iteratively. The final selected regions are listed in Table I. All these features are ultimately concatenated to form a long feature vector. Note, in this paper, the samples we consider for verification are the absolute difference between two subjects rather than the single image or the corresponding feature. Specifically, the sample of child–old parent pair x_i and child–young parent pair x_j can be written as $x_i = |F_O - F_C|$ and $x_j = |F_Y - F_C|$, where F_O, F_Y, and F_C are features of old–parent, young–parent, and child, respectively.

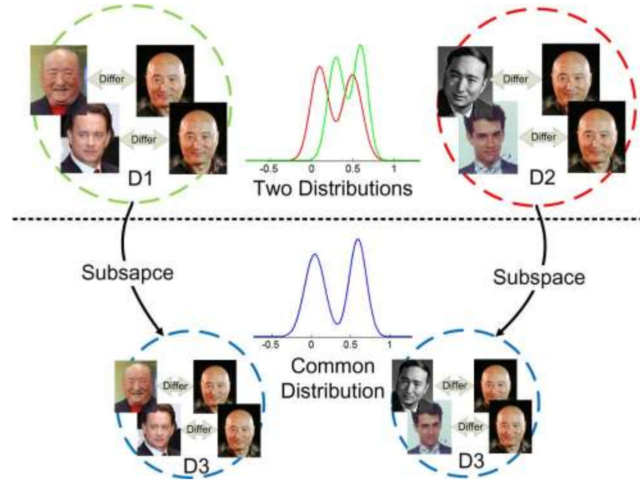


Fig. 6. Transfer subspace learning for kinship verification. D1, D2, and D3 represent target domain, source domain, and learned subspace. The source and target domains have different distributions before they are projected to the learned subspace. With transfer learning, features from both source and target domains now share a common distribution. Note the first and second image pairs in D1, D2, and D3 represent true and false child–parent pairs, respectively.

C. Transfer Subspace Learning

In our framework, we attempt to find a subspace where distribution similarity of two different data sets, i.e., child–old parent and child–young parent pairs, is maximized while in the subspace they still can be well separated in terms of kinship verification. Essentially, our approach is different from [42] in that ours takes advantage of the intermediate set (young parents) and prevents the failure of transfer. Moreover, instead of directly modeling three distributions similar to [27], we use two distributions by pairwise differences of three data sets, which leads to a simple but efficient model.

In terms of transfer learning [32], method in [27] considers children, young parents and old parents as the source, intermediate and target domain, respectively. However, the differences of features between children and parents are more discriminative when we conduct two-class classification. So in this paper, we impose transfer learning on child–old parent and child–young parent pairs and therefore they become the source and target domain of the new problem. Our objective can be formulated as, finding an appropriate subspace where the low-dimensional representation of child–young parent feature difference and that of child–old parent are still discriminative and share the same distribution. Fig. 6 illustrates the idea of transfer subspace learning proposed in this paper which can be mathematically formulated as

$$W = \arg\min_{W} \{F(W) + \lambda D_W(P_L \| P_U)\} \tag{2}$$

where $F(W)$ is a general subspace learning objective function, such as PCA [4], LDA [6], LPP [7], NPE [8], MFA [9], and DLA [10], and P_L and P_U represent the distribution of the source and target samples respectively. $D_W(P_L \| P_U)$ is the Bregman divergence-based regularization that measures the distance between two different distributions in the projected subspace W. λ is the weight for the regularization. This optimization problem can be solved iteratively by the gradient descent method as follows:

$$W_{k+1} = W_k - \eta(k)\left(\frac{\partial F(W)}{\partial W} + \lambda \sum_{i=1}^{l+u} \frac{\partial D_W(P_L \| P_U)}{\partial \vec{y}_i} \frac{\partial \vec{y}_i}{\partial W}\right) \tag{3}$$

where $\eta(k)$ is the learning rate, $y = W^T x$ is the low-dimensional representation of the sample, and l and u are the numbers of samples in the distribution of P_L and P_U. Particularly, the second term with the gradient in (3) can be obtained based on a discriminant subspace learning method, e.g., discriminative locality alignment (DLA) [10]

$$\frac{\partial F(W)}{\partial W} = (XLX^T + (XLX^T)^T)W \tag{4}$$

where L encapsulates the local geometry and the discriminative information and X is the matrix of input samples of all the training images. Moreover, estimations of P_L and P_U of (2) in the projected subspace can be achieved by kernel density estimation (KDE) [42]. Therefore, projection matrix W can be updated iteratively until it is optimized. By introducing the intermediate set (young parents) and shrinking the discrepancy between feature difference of child–young and that of child–old parent, finally, the child–old parent verification problem becomes more tractable. The experimental results in Section IV will demonstrate that the verification accuracy with the intermediate data set is better than that with direct matching from children to old parents.

D. Kinship Recognition With Context

In this section, we consider the task of kinship verification in a photo as a joint tagging problem [1] to further improve the accuracy. In the real-world people-centric images, relationships are not independent anymore. For example, kin relationships often accompany couple relationships in the same image. In addition, labels such as age and gender of two people are also helpful in determining their relations. For example, people with age gap around 20 years have high probability to be child–parent pairs. Different from [1] that considers labels of each node and relations between labels simultaneously, we instead first predict labels of each individual and then seek possible relations. In this way, complex constrains in one step now are decomposed into two separate steps, which improve computational efficiency. For simplicity, we will assume that the image has already been parsed into a discrete set of face regions via an existing face detection algorithm. We further assume that each of the detected faces is associated with a discrete set of allowable labels, i.e., gender and age. Our goal is to infer a joint tagging matrix $R = r_{ij}$ from the entire feature space X and the single label vector for each person $Y = (y_1, y_2, \ldots, y_n)$. This joint tagging progress can be expressed as

$$h(X, Y) = \arg\max_{R} f(X, Y, R) \tag{5}$$

where $f(X, Y, R)$ is the total score over all individuals of this joint annotation work. Suppose there are three relations we will consider, i.e., kinship, couple, and others. So the joint annotation function $f(X, Y, R)$ can be written as

$$f(X, Y, R) = \sum_{i \neq j} \beta_{r_{ij}} \phi_{r_{ij}}(r_{ij} | y_i, y_j, X) \tag{6}$$

where r_{ij} is an element in the relation matrix R and it could be equal to 1, 2, or 3, which are corresponding to kin, couple, and other relationships. $\beta_{r_{ij}}$ is the prior for each of three relations and can be readily computed by the training data. For each relation $\phi_{r_{ij}}(r_{ij} | y_i, y_j, X)$, we will use the following linear regression model to formulate:

$$\phi_{r_{ij}}(r_{ij} | y_i, y_j, X) = w_{r_{ij}}^T z_{ij} \tag{7}$$

where $w_{r_{ij}}$ is the weight vector for the feature vector z_{ij}. For different relations, there should be three different weight vectors w_1, w_2 and w_3. The 4-D feature vector $z_{ij} = (z_g, z_a, z_d, z_k)^T$ comprises four parts, i.e., gender relation, age difference, relative distance, and kinship score, by which contextual information embedded in individuals y_i and y_j can be utilized. Details of each dimension are defined as follows.

- **Gender relation**: Gender is highly related to the couple relationship. People with different gender in the same age group have higher probability of being couple. If individuals y_i and y_j are with the same gender, then $z_g = 1$; otherwise 0.
- **Age difference**: Age difference also affects the relationship between two people. Generally, parents and children should have an age gap from 20 to 30 with high probabilities. Therefore, we use age difference between individuals y_i and y_j as z_a.
- **Relative distance**: Distance between two people measures the intimacy. We use the center of two eyes as the location of an individual and calculate the Euclidean distance between two individuals. It is normalized by the average over all pairwise distances of all individuals.
- **Kinship score**: Kinship score can be calculated by the proposed transfer subspace learning. For K-NN method, we will count the positive and negative samples and calculate their ratio over K neighbors. The percentage of positive samples will be the score for z_k.

The solution of (5) is not trivial. We can traverse all the possible relationship combinations and achieve the largest score, since in our experiment, the number of people in one image will not prohibitively large.

IV. Experiments

Before following experiments, we conduct face and fiducial points detection on all images in the UB KinFace and FamilyFace database to obtain cropped face images with the fixed size of 127×100. Then we register faces according to corresponding points detected by the last step using an affine transform. Several image preprocessing techniques, e.g., histogram equalization, are implemented to mitigate image quality factors. After that, discriminative features are extracted based on the method proposed in Section III. A series of experiments are specifically designed to verify the proposed assumption, the transfer subspace learning method and the context based verification. We also compare the performance by human judges with that of the method proposed in this paper to present a systematic evaluation. Since most young parent images are captured with outdated equipment, the black and white images are common in this category. Therefore, for fair comparisons with other approaches that do not involve young–parent images, we run all the experiments on gray-scale images. The designed experiments in this section were carried out on an Intel Core 2 Duo CPU E7300 2.66 GHz and 3 GB memory, and all algorithms were implemented by Matlab 7.0.

A. Experiments on UB KinFace Database

1) Kinship Classification: First, a kinship classification experiment is performed. Here in terms of biometrics, classification means finding a proper identity for the query. Specifically, in this section, our aim is that given a child's facial image, we will seek and return his/her parent's image, either young or old. In this process, children images are used as queries while young parents and old parents images are used as gallery. Euclidean distance metric and nearest-neighbor classifier are adopted for this task and results are shown in Fig. 7(a), from which we can see that young parents are more similar to their children than old parents based on the local Gabor features proposed in this paper.

2) Kinship Verification: We conduct kinship verification to further prove our hypothesis: given two images of faces, determine if they are the true child–parent pair. In these experiments, rather than direct comparisons between children and their parents, feature discrepancy that measures the difference between the child and the parent is used. For the purpose of training and testing, we use 200 true child–old parent pairs and 200 false child–old parent pairs. Both anthropometric model and proposed method are evaluated. In order to classify the image pairs into the true and false child–parent pairs, we use

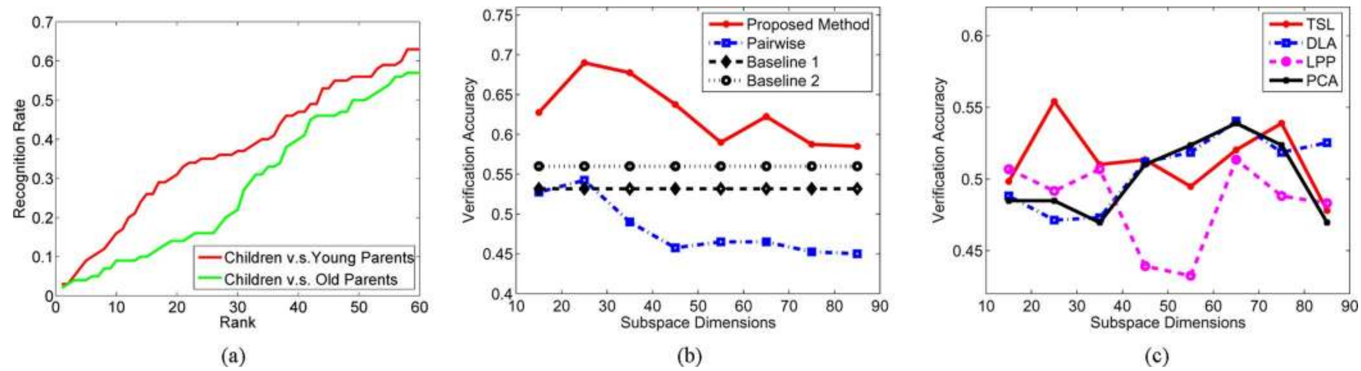


Fig. 7. Left: Cumulative match characteristic (CMC) for kinship classification. "Rank" indicates the size (number) of candidate list for comparison. Middle: Pairwise and transfer subspace learning over different subspace dimensions on kinship verification by leave-one-out strategy. "Pairwise" represents the direct matching with PCA and "Baseline 1" and "Baseline 2" the human performances mentioned in Section V. Right: Kinship verification by PCA, LPP, DLA and the proposed transfer subspace learning (TSL) method over different subspace dimensions in the FamilyFace database.

TABLE II
VERIFICATION RESULTS BY FIVE-FOLD CROSS VALIDATION. "C VERSUS. Y" AND "C VERSUS. O" MEAN CHILD–YOUNG PARENT AND CHILD–OLD PARENT VERIFICATION, RESPECTIVELY. "RAW" AND "STRUCTURE" INDICATE THE RAW IMAGE FEATURE AND STRUCTURE FEATURE, RESPECTIVELY

Verification 5 fold	C vs. Y (Raw)	C vs. O (Raw)	C vs. Y (Structure)	C vs. O (Structure)
Average	**53.89%**	50.56%	**56.67%**	53.33%

TABLE III
VERIFICATION COMPARISONS BETWEEN PAIRWISE AND THE PROPOSED TRANSFER SUBSPACE LEARNING (TSL) METHODS WITH DIFFERENT FEATURES

Feature	Local Gabor (ours)	LBP [12]	Structure [47]	Raw
Pairwise	50.25%	50.00%	49.00%	48.50%
TSL	**56.50%**	52.00%	47.25%	49.50%

Euclidean distance and KNN classifier with five-fold cross validation where 40 positive sample pairs and 40 negative sample pairs are used as test set at each round. Particularly, the positive samples are the true pairs containing children and parents and negative ones are children with randomly selected parents who are not their true parents. Results are shown in Table II.

Raw images and structure features based experimental results with five-fold cross validation in Table II still support our hypothesis. The kinship verification rate based on young parents is about three percent higher than that of based on old parents. Considering the poor quality of Internet-collected "wild" images of young parents (It is impractical to acquire high quality images since the digital camera was not popular back to that time), the improvements are significant enough to validate our hypothesis.

3) Kinship Verification With Transfer Learning: In the training phase of the transfer subspace learning based verification, the inputs are 320×2 samples (for five-fold cross validation). Half of inputs (320 samples) are child–old parent pairs while the rest of them are child–young parent pairs. As DLA [10] is a supervised method, both 320 child–old parent pairs and child–young parent pairs consist of 160 true kinship pairs and 160 false ones. For the first part of (3), we initialize W with the one generated only by the source, namely, the true and false child–young parent pairs. As to the second term in (3), we calculate distributions of source (child–young parent pairs, including both true and false ones) and target (child–old parent pairs, including both true and false ones) by KDE [42]. In the test phase, we use 80 child–old parent (40 positive and 40 negative for five-fold cross validation) samples for verification experiment. All image pairs, either training or testing, are projected into the optimal subspace before KNN classification with Euclidean distance. Additionally, we compare the performance of the proposed feature extraction method with anthropometric model. The configuration is the same as the experiments above except including young parent samples. Fig. 7(b) shows the results of pairwise and transfer subspace learning produced by leave-one-out strategy over different subspace dimensions. The average training time for pairwise and transfer subspace learning is 2.84 and 154 seconds respectively given that the subspace dimension is 25. The average time for verification is 4.49 milliseconds. Table III lists several results of kinship verification by different methods and features by five-fold cross validation.

4) Kinship Verification by Human: Few works have been done to evaluate human performance on child–parent kinship verification though individual identification and sibling relationship verifications have been reported in [15] and [30]. To extensively evaluate our method and probe the feasible learning strategy by which human being perceive kin relationship, we proceed with two groups of human tests on kinship verification. The first group has 40 training samples (20 true child–old parent pairs as well as 20 false child–old parent pairs) and 40 test samples (20 true child–old parent pairs as well as 20 false child–old parent pairs, not overlapped with the training ones). In the second group, however, 20 extra corresponding true child–young parent pairs are added as supplement of training samples and all other sets are identical with the former ones.

The two experiments mentioned above aim at simulating the process of pairwise and transfer subspace learning based verification. Twenty people participate in these experiments and average performances are 53.17% and 56.00% for the first and second experiment which are illustrated in Fig. 7(b) as two baselines. Compared with the best performance of pairwise and transfer subspace learning based verification, humans perform better than most of them with the exception

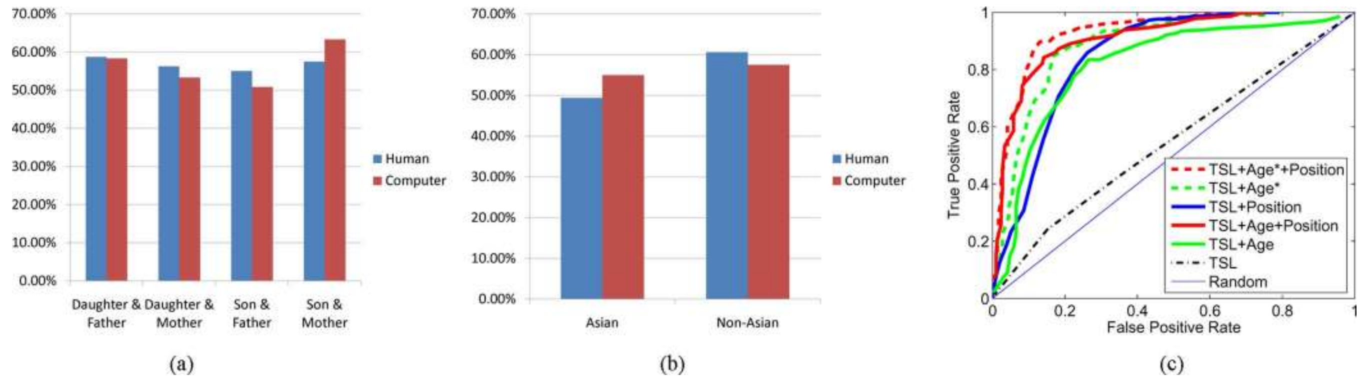


Fig. 8. Kinship verification on four kinds of kinship relations (Left) and two kinds of races (Middle) based on the FamilyFace database. ROC curve of pairwise and context based kinship verification (Right). Here we use random guess as the baseline. "TSL" is the proposed transfer subspace learning method and labels like "Age" and "Position" are added one by one. Note "Age*" uses ground truth labels while "Age" uses predicted labels via method in [20]. Specifically, we do not add gender into this experiment in that it is weakly relevant to the direct kinship inference.

of the proposed method, as Fig. 7(b) illustrated. In fact, even without training samples, humans can perform well due to their sufficient knowledge. This is mostly accumulated when people grow up and helps people perform recognition. Therefore, human being are not as sensitive as machines against new training samples. Given more training data, e.g., child–young parent pairs, however, machines can learn and acquire more discriminative information.

B. Experiments on FamilyFace Database

There are two purposes for the experiments of this section. One is to investigate the impact of gender and race on kinship, and the other is to evaluate the role of contexts and semantics in family albums. For the first purpose, we use the UB KinFace Ver2.0, FamilyFace and images downloaded from the Web as the training and testing data. The entire UB KinFace is used for training while 296 true and 296 false child–parent pairs from the FamilyFace and the Web are used for testing. Given a family photo, we first detect faces and extract features. Then the difference of each pair of face features is projected into the subspace learned from the training set followed by KNN classification. Experimental results are shown in Fig. 7(c). To further explore the impact of the four typical kin relationships on verification accuracy, we select 120 (60 positive and 60 negative) samples with the same kin relationship and repeat this for the four relationships. As can be seen from Fig. 8(a), verification rates on "daughter–father" and "son–mother" are higher than those on "daughter–mother" and "son–father". It suggests that gender affects the determination of kinship to some extent. In addition, race is taken into consideration since the appearance diverges over different races. Fig. 8(b) illustrates the verification results on two race groups, i.e., Asian and non-Asian.

In paralleled with machine based experiments on the FamilyFace database, we conduct another one accomplished by human judges. In these experiments, 20 participants determine the authenticity of kin relationships of 32 pairs which are randomly selected from FamilyFace database. Therefore, there are 8 pairs for each relation and half of them are positive samples. The experimental results show that the average accuracy by human on "daughter–father" relationship is 58.75%, "daughter–mother" 56.25%, "son–father" 55.00%, and "son–mother" 57.50%, respectively. We also separate the data set into Asian and non-Asian subsets and evaluate the human performance on kinship verification. The results are 49.41% and 60.59%, respectively. It is interesting to notice that the verification accuracy for non-Asian is higher than that of Asian. This could be explained by the fact that most non-Asians have more distinctive facial features.

TABLE IV
"PROPOSED METHOD" MEANS METHOD BASED ON (6), "LINEAR MODEL" MEANS CONTEXT BASED METHOD ONLY CONSIDERING "KINSHIP" AND "TSL" MEANS THE TRANSFER SUBSPACE LEARNING METHOD ONLY

Method	True positive	False positive
Proposed method	79.66%	10.68%
Linear model	74.57%	15.53%
TSL	63.56%	55.34%

Contextual information based experiments use both FamilyFace database (119 images) and images downloaded from the Web (100 images), e.g., Facebook, which contain friend and colleague relationships as well. Experiment in this part is divided into two steps. The first step only uses the regression model in (7) and only kinship is considered which aims at evaluating the effectiveness of "gender relation," "age difference," "relative distance," and "kinship score." The gender and age estimation methods are from [20]. Ten-fold cross validation is used and results are shown in Fig. 8(c). As we can see, the more contextual information (age, position) is added, the higher ROC score is obtained. Moreover, we show the impact of unreliable estimation in Fig. 8(c). Since we use the method in [20] to estimate age, estimation error is inevitable. In Fig. 8(c), the green dot dash line and green solid line represent the performance of kinship verification with ground truth age label and estimated value, respectively. The divergence between these two lines in Fig. 8(c) shows the aforementioned impact of unreliable age estimation. The second step of the experiment is to incorporate the couple and other relations to restrict the relation function in (7), therefore leading to an even better result. We use approximately half of the images for training and the other half for testing. Then with (6), we obtain the results shown in Table IV. The average time for relationship recognition is 24.8 seconds. Note that subspace dimensions for "TSL" method is fixed at 25 in Fig. 8 and Table IV.

C. Discussion

From the experimental results above, it can be concluded that the appearance similarity gap is so large that the task of child–parent classification is very challenging. Nevertheless, for child–parent verification, the proposed approach based on transfer subspace learning outperforms pairwise kinship verification, validating that our approach of narrowing the similarity gap between the child and his/her parent in terms of appearance is effective. In addition, we find that the proposed local feature selection strategy is more discriminative than anthropometric models due to a lack of discriminative information in the latter (six dimensions). This is consistent with the intuition that people tend to recognize the kin relationship by comparing facial features, such as eyes, nose, and mouth.

Compared with the result in [25], some factors prevent a better experimental result. First, variations in pose, lighting and aging of faces in our data set are more significant. Second, color information is not taken advantage of in our experiment due to the poor quality of young parents images. As a matter of fact, family members always have similar hair, eyes and skin colors. In addition, the volume and structure of two databases are different. The dataset used in [25] has 150 groups of 300 images while ours has 200 groups of 600 images and considers different race distributions as well. Finally, though a direct comparison with [25] is not practical, we have already proved that transfer subspace learning based method is better than mere pairwise verification which is utilized by [25] in the discriminant phase.

Moreover, contextual information is useful and results are significantly improved due to more available semantic constraints. However, we should point out that both gender and age estimations are imperfect in practice. Improvements of their solutions will certainly benefit kinship verification, as illustrated in Fig. 8(c). Furthermore, in this paper, the family album does not contain too many people and relationships we consider are only kinship, couple and others. When there are more people in images, more complex relationships should be considered, e.g., siblings, inferior/superior relationships. In addition, some differences exist between our testing results and those in [25], for instance, impacts of gender and race. This is due to the limited capacity and different bias in each database. In any cases, it proves that both gender and race should be carefully considered for kinship verification.

V. Conclusion

In this paper, we investigate the problem of kinship verification through face images as well as the impact of contexts and semantics. First, the UB KinFace Ver2.0 was collected from the Web. Second, we propose a transfer subspace learning method using young parents as an intermediate set whose distribution is close to both the source and target sets. Through this learning process, the large similarity gap between distributions is reduced and child–old parent verification problem becomes more tractable. Third, we combine the proposed transfer learning approach with contextual information in family albums to further improve verification accuracy. Experimental results demonstrate our hypothesis on the role of young parents is valid and transfer learning can take advantage of it to enhance the verification accuracy. In addition, we prove that contextual information can reasonably improve the kinship verification accuracy via tests on the FamilyFace database.

Acknowledgment

This work was performed at SUNY–Buffalo when S. Xia was a Visiting Scholar there. S. Xia and M. Shao are equal contributors to this paper.

References

[1] Z. Stone, T. Zickler, and T. Darrell, “Toward large-scale face recognition using social network context,” *Proc. IEEE*, vol. 98, no. 8, pp. 1408–1415, Aug. 2010.

[2] P. Viola and M. Jones, “Robust real-time face detection,” *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.

[3] G. Huang, V. Jain, and E. Learned-Miller, “Unsupervised joint alignment of complex images,” in *IEEE Int. Conf. Comput. Vis.*, Rio de Janeiro, Brazil, 2007.

[4] M. Turk and A. Pentland, “Eigenfaces for recognition,” *J. Cognitive Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

[5] M. Bartlett, J. Movellan, and T. Sejnowski, “Face recognition by independent component analysis,” *IEEE Trans. Neural Netw.*, vol. 13, no. 6, pp. 1450–1464, Nov. 2002.

[6] P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 2002.

[7] X. He and P. Niyogi, “Locality preserving projections,” in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2004.

[8] X. He, D. Cai, S. Yan, and H. Zhang, “Neighborhood preserving embedding,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2005, pp. 1208–1213.

[9] 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.*, vol. 29, no. 1, pp. 40–51, Jan. 2007.

[10] T. Zhang, D. Tao, and J. Yang, “Discriminative locality alignment,” in *Proc. Eur. Conf. Comput. Vis.*, 2008, pp. 725–738.

[11] Y. Adini, Y. Moses, and S. Ullman, “Face recognition: The problem of compensating for changes in illumination direction,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 721–732, Jul. 2002.

[12] T. Ahonen *et al.*, “Face description with local binary patterns: Application to face recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

[13] Z. Cao, Q. Yin, X. Tang, and J. Sun, “Face recognition with learning-based descriptor,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 2707–2714.

[14] M. Guillaumin, J. Verbeek, and C. Schmid, “Is that you? Metric learning approaches for face identification,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 498–505.

[15] N. Kumar, A. Berg, P. Belhumeur, and S. Nayar, “Attribute and simile classifiers for face verification,” in *Proc. IEEE Int. Conf. Comput. Vis.*, 2009, pp. 365–372.

[16] K. Jia and S. Gong, “Generalized face super-resolution,” *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 873–886, Jun. 2008.

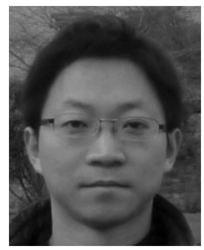
[17] C. Liu, H. Shum, and W. Freeman, “Face hallucination: Theory and practice,” *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 115–134, 2007.

[18] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, “Labeled faces in the wild: A database for studying face recognition in unconstrained environments,” Univ. Massachusetts, Amherst, MA, Tech. Rep. 07-49, 2007.

[19] M. Naaman, R. Yeh, H. Garcia-Molina, and A. Paepcke, “Leveraging context to resolve identity in photo albums,” in *Proc. ACM/IEEE-CS Joint Conf. Digital Libraries*, 2005, pp. 178–187.

[20] A. Gallagher and T. Chen, “Understanding images of groups of people,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 256–263.

[21] G. Wang, A. Gallagher, J. Luo, and D. Forsyth, "Seeing people in social context: Recognizing people and social relationships," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 169–182.
[22] J. Yagnik and A. Islam, "Learning people annotation from the web via consistency learning," in *Proc. Int. Workshop Multimedia Inf. Retrieval*, 2007, pp. 285–290.
[23] T. L. Berg, E. C. Berg, J. Edwards, M. Maire, R. White, Y. whye Teh, E. Learned-miller, and D. A. Forsyth, "Names and faces in the news," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2004, pp. 848–854.
[24] A. Gallagher and T. Chen, "Clothing cosegmentation for recognizing people," presented at the IEEE Conf. Comput. Vis. Pattern Recognit., Anchorage, AK, Jun. 2008.
[25] R. Fang, K. Tang, N. Snavely, and T. Chen, "Towards computational models of kinship verification," in *Proc. IEEE Int. Conf. Image Process.*, 2010, pp. 1577–1580.
[26] S. Ng, C. Lai, M. Tsai, R. Fang, and T. Chen, "Automatic kinship recognition using social contextual information," presented at the Bits on Our Minds, Itchaca, NY, 2011.
[27] S. Xia, M. Shao, and Y. Fu, "Kinship verification through transfer learning," in *Proc. Int. Joint Conf. Artificial Intell.*, 2011, pp. 2539–2544.
[28] M. Shao, S. Xia, and Y. Fu, "Genealogical face recognition based on UB kinface database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (Workshop Biometrics)*, 2011, pp. 65–70.
[29] A. Alvergne, C. Faurie, and M. Raymond, "Differential facial resemblance of young children to their parents: Who do children look like more?," *Evol. Human Behavior*, vol. 28, no. 2, pp. 135–144, 2007.
[30] M. D. Martello and L. Maloney, "Where are kin recognition signals in the human face?," *J. Vis.*, vol. 6, no. 12, pp. 1356–1366, 2006.
[31] L. DeBruine, F. Smith, B. Jones, S. C. Roberts, M. Petrie, and T. Spector, "Kin recognition signals in adult faces," *Vis. Res.*, vol. 49, no. 1, pp. 38–43, 2009.
[32] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowledge Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
[33] W. Dai, Q. Yang, G. Xue, and Y. Yu, "Boosting for transfer learning," in *Proc. Int. Conf. Machine Learning*, 2007, pp. 193–200.
[34] R. Raina, A. Battle, H. Lee, B. Packer, and A. Ng, "Self-taught learning: Transfer learning from unlabeled data," in *Proc. Int. Conf. Mach. Learning*, 2007, pp. 759–766.
[35] J. Jiang and C. Zhai, "Instance weighting for domain adaptation in NLP," in *Proc. Annu. Meeting–Assoc. Comput. Ling.*, 2007, vol. 45, pp. 264–264.
[36] A. Evgeniou and M. Pontil, "Multi-task feature learning," in *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2007, vol. 19.
[37] N. Lawrence and J. Platt, "Learning to learn with the informative vector machine," in *Proc. Int. Conf. Machine Learning*, 2004, pp. 512–519.
[38] L. Mihalkova, T. Huynh, and R. Mooney, "Mapping and revising Markov logic networks for transfer learning," in *Proc. AAAI*, 2007, vol. 22, no. 1, pp. 608–614.
[39] A. Arnold, R. Nallapati, and W. Cohen, "A comparative study of methods for transductive transfer learning," in *Proc. Int. Conf. Data Mining (Workshops)*, 2007, pp. 77–82.
[40] W. Dai, G.-R. Xue, Q. Yang, and Y. Yu, "Transferring naive Bayes classifiers for text classification," in *Proc. AAAI*, 2007, pp. 540–545.
[41] W. Dai, G. Xue, Q. Yang, and Y. Yu, "Co-clustering based classification for out-of-domain documents," in *Proc. ACM SIGKDD Conf. Knowl. Discovery Data Mining*, 2007, pp. 210–219.
[42] S. Si, D. Tao, and B. Geng, "Bregman divergence-based regularization for transfer subspace learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 7, pp. 929–942, Jul. 2010.
[43] Y. Su, Y. Fu, Q. Tian, and X. Gao, "Cross-database age estimation based on transfer learning," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process.*, 2010, pp. 1270–1273.
[44] A. Gallagher and T. Chen, "Using group prior to identify people in consumer images," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2007, pp. 1–8.
[45] P. Singla, H. Kautz, J. Luo, and A. Gallagher, "Discovery of social relationships in consumer photo collections using Markov logic," presented at the IEEE Conf. Comput. Vis. Pattern Recognit., Washington, DC, 2008.
[46] J. Wong and S. Cho, "A face emotion tree structure representation with probabilistic recursive neural network modeling," *Neural Comput. Appl.*, vol. 19, no. 1, pp. 33–54, 2010.
[47] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 387–394.
[48] T. Chen, W. Yin, X. Zhou, D. Comaniciu, and T. Huang, "Total variation models for variable lighting face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1519–1524, Sep. 2006.
[49] A. Jain and S. Li, *Handbook of Face Recognition*. New York: Springer, 2005.



Siyu Xia received the B.E. and M.S. degrees in automation engineering from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2000 and 2003, respectively, and the Ph.D. degree in pattern recognition and intelligence system from Southeast University, Nanjing, China, in 2006.

From 2010 to 2011, he was a Visiting Scholar in the Department of Computer Science and Engineering, State University of New York at Buffalo. He currently works at Southeast University, Nanjing, China. His research interests include object detection, applied machine learning, social media analysis, and intelligent vision system.

Ming Shao received the B.E. degree in computer science, the B.S. degree in applied mathematics, and the M.E. degree in computer science and engineering from Beihang University, Beijing, China, in 2006, 2007, and 2009, respectively. He is currently working toward the Ph.D. degree in computer science and engineering at the State University of New York at Buffalo.

His current research interests include sparse coding, low-rank representation, graph theory, multi-label learning, and applied machine learning on social media.

Jiebo Luo (S'93–M'96–SM'99–F'09) received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology of China (USTC), Hefei, China, in 1989 and 1992, respectively; and the Ph.D. degree in electrical and computer engineering from the University of Rochester, Rochester, NY, in 1995.

He was a Senior Principal Scientist with the Kodak Research Laboratories, Rochester, NY, before joining the Computer Science Department, University of Rochester, in 2011. His research interests include signal and image processing, machine learning, computer vision, social media data mining, and medical imaging. He has authored over 180 technical papers and holds over 60 U.S. patents.

Dr. Luo is a Fellow of the SPIE and IAPR. He has been actively involved in numerous technical conferences, including serving as the General Chair of ACM CIVR 2008; program Co-Chair of IEEE CVPR 2012 and ACM Multimedia 2010; area Chair of IEEE ICASSP 2009–2012, ICIP 2008–2012, CVPR 2008, and ICCV 2011. He has served on the editorial boards of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, *Pattern Recognition, Machine Vision and Applications*, and the *Journal of Electronic Imaging*.

Yun Fu (S'07–M'08–SM'11) received the B.Eng. degree in information engineering and the M.Eng. degree in pattern recognition and intelligence systems from Xi'an Jiaotong University (XJTU), China, in 2001 and 2004, respectively; and the M.S. degree in statistics and the Ph.D. degree in electrical and computer engineering from the University of Illinois at Urbana–Champaign (UIUC), in 2007 and 2008, respectively.

He was a Research Intern with Mitsubishi Electric Research Laboratories, Cambridge, MA, in summer 2005; with Multimedia Research Lab of Motorola Labs, Schaumburg, IL, in summer 2006. He jointed BBN Technologies, Cambridge, MA, as a Scientist in 2008. He holds a part-time Lecturer position at the Department of Computer Science, Tufts University, Medford, MA, in the spring of 2009. He joined the Department of Computer Science and Engineering, SUNY at Buffalo as an Assistant Professor in 2010. His research interests include applied machine learning, social media analytics, human-centered computing, pattern recognition, and intelligent vision system.

Dr. Fu is a life member of ACM, SPIE, AAAI, and Institute of Mathematical Statistics (IMS), and 2007–2008 Beckman Graduate Fellow. He is the Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (TCSVT). He is the recipient of the 2002 Rockwell Automation Master of Science Award, Edison Cups of the 2002 GE Fund Edison Cup Technology Innovation Competition, the 2003 Hewlett-Packard (HP) Silver Medal and Science Scholarship, the 2007 Chinese Government Award for Outstanding Self-Financed Students Abroad, the IEEE ICIP 2007 Best Paper Award, the 2007–2008 Beckman Graduate Fellowship, the 2008 M. E. Van Valkenburg Graduate Research Award, the ITESOFT Best Paper Award of 2010 IAPR International Conferences on the Frontiers of Handwriting Recognition (ICFHR), the 2010 Google Faculty Research Award, the Best Paper Award for the IEEE ICDM 2011 Workshop on Large Scale Visual Analytics, IEEE ICME 2011 Quality Reviewer, and the 2011 IC Postdoctoral Research Fellowship Award.