Neurocomputing 214 (2016) 350–357

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Mixed bi-subject kinship verification via multi-view multi-task learning

Xiaoqian Qin [a,b,c], Xiaoyang Tan [a,b,*], Songcan Chen [a,b]

[a] College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
[b] Collaborative Innovation Center of Novel Software Technology and Industrialization, Nanjing 210016, China
[c] Huaiyin Normal University, Huaian 223300, China

ARTICLE INFO

Article history:
Received 1 December 2015
Received in revised form
7 June 2016
Accepted 9 June 2016
Communicated by Dr Jiwen Lu
Available online 16 June 2016

Keywords:
Mixed-kinship verification
Multi-task learning
Multi-view learning

ABSTRACT

Bi-subject kinship verification addresses the problem of verifying whether there exists some kind of kin relationship (i.e., father–son, father-daughter, mother–son and mother–daughter) between a pair of parent–child subjects based purely on their visual appearance. The task is challenging due to the involvement of two different subjects possibly with different genders and ages. In addition, collecting sufficient training samples for each type of kinship is difficult. In this work, we present a novel method to address these issues by considering each type of kin relation verification as one task and learning them at one time in the framework of multi-task learning, by sharing feature sets and useful structures among the tasks.

Particularly our contributions are three folds: first, we introduce a new type of learning problem, called mixed bi-subject kinship verification, to the topic of bi-subject kinship verification: instead of simply verifying whether some fixed kinship relationship (e.g., mother–son) can be established for a given pair of parent–child images, we try to figure out whether any type of the four kinship relations can be established according to the visual features of the image pair, with no need to know the genders of the subjects to be verified beforehand. Second, we propose a novel multi-task learning method to address this problem with two transformation matrices – one is shared amongst all the tasks and the other is unique to each task. Both matrices are simultaneously learned in a joint framework, which enables our algorithm to utilize the common knowledge of the four tasks. Third, we propose a multi-view multi-task learning(MMTL) method to perform multiple feature fusion to improve the mixed bi-subject kinship verification performance. Extensive experiments on the large scale KinFaceW kinship database demonstrate the feasibility and effectiveness of the proposed algorithm.

1. Introduction

The aim of kinship verification in the field of computer vision is to predict whether a few given subjects have certain kind of kin relation purely through their visual appearance. Applications of this include face image retrieval [6,22,37], face annotation [28,57], face images organization, face recognition [36,40], social media analysis [45,55], children adoptions [18], and so on.

However, the problem of kinship verification is challenging, since it involves multiple subjects possibly with different genders and ages, rather than multiple images of the same subject as in the traditional face verification. On the other hand, the researches in the field of human visual signal processing [7,9] have provided strong evidence that facial appearance is a useful cue for genetic similarity – children look more similar to their parents than other adults of the same gender. There exists some progress in finding distinguishable cues from facial appearance [11,15,42,43,27,47,53]. While most of them focus on bi-subject kinship verification, i.e. father–son, father–daughter, mother–son and mother–daughter, Qin et al. [32] tackle the problem of tri-subject kinship verification, i.e., father/mother–son and father/mother–daughter. However, all these works assume that the genders of the pair of subjects be known before doing verification. In other words, they can only verify one kind of kinship at one time.

In this work, we introduce a new type of learning problem, called mixed bi-subject kinship verification. That is, instead of simply verifying whether one fixed kinship relationship (e.g., mother–son) can be established for a given pair of parent–child images, we try to figure out whether any type of the four kinship relations can be established according to the visual features of the image pair, with no need to know the genders of the subjects to be

* Corresponding author at: College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China.
E-mail address: x.tan@nuaa.edu.cn (X. Tan).

http://dx.doi.org/10.1016/j.neucom.2016.06.027
0925-2312/

verified beforehand. This is not only a step toward more practical kinship verification, but also imposes interesting challenges onto this topic – without knowing the genders of subjects, the problem of bi-subject kinship verification can not be simply treated as a binary classification problem. In addition, existing works show that the gender plays a significant role in kinship verification [11,27,32,47,53]. And earlier psychological research results also indicate that the kin recognition signal is different under different gender [31].

According to the studies of human genetics, we know that some facial features, such as eye or hair color or dimples have been proposed to be specifically linked to paternity confidence [34]. Hair, skin and eye colors are highly heritable and visible traits in humans [38]. Thus, we think that different kinship relations may have some common genetic characteristics. Motivated by these researches, we propose a novel multi-task learning method to address this problem with two transformation matrices – one is shared amongst all the tasks and the other is unique to each task. Both matrices are simultaneously learned in a joint framework, which enables our algorithm to utilize the common knowledge of the four tasks. Further, compared to single task learning, the advantage of multi-task learning is usually more visible when we only have a few training examples per task. We expect that even if only a few training samples are available for each task, the use of multi-task learning can compensate discriminative information by transferring from one task to the others. To make better use of multi-feature descriptors, we propose a multi-view multi-task learning (MMTL) method to perform multiple feature fusion to improve the mixed bi-subject kinship verification performance. Extensive experiments on the large KinFaceW kinship database demonstrate the feasibility and effectiveness of the proposed algorithm. The contributions of this work are summarized as follows:

(1) A new type of learning problem called mixed bi-subject kinship verification is proposed to the topic of bi-subject kinship verification. We try to figure out whether any type of the four kinship relations can be established according to the visual features of the image pair, with no need to know the genders of the subjects to be verified beforehand.
(2) A novel multi-task learning method is proposed to address our proposed problem with two transformation matrices – one is shared amongst all the tasks and the other is unique to each task. Both matrices are simultaneously learned in a joint framework, which enables our algorithm to utilize the common knowledge of the four tasks.
(3) A multi-view multi-task learning method is developed to perform multiple feature fusion to improve the mixed bi-subject kinship verification performance.

In what follows, we briefly review some of the related work in Section 2 and detail our proposed method in Section 3. Experimental results are given in Section 4. We conclude this paper in Section 5.

2. Related work

Existing works on kinship learning can be roughly be divided into two categories: kinship classification and verification. Fang et al. [14] tackle the former one, i.e., given a query face image, asking which family it belongs to. They do this with a minimum sparse reconstruction method.

Most of the works are about kinship verification. In an early attempt, Fang et al. [15] use various features including the skin, hair and eye color, facial structure measures and local/holistic texture. Later, researchers evaluate various types of feature descriptors including DAISY [18], Gradient Orientation Pyramid (GGOP) [55], Gated Autoencoder [10], Self Similarity Representation (SSR) [23], prototype-based discriminative feature learning (PDFL) method [48], a spatial pyramid learning-based feature descriptor Zhou et al. [54], dense stereo matching [36] and attributes [43]. Dibeklioglu et al. [11] propose to describe facial appearance over smile expression.

In [27], the authors show that combining several types of middle-level features is useful. For that purpose, they introduce a multi-view neighborhood repulsed metric learning method (MNRML) by learning a distance metric. Yan et al. [47] and Hu et al. [20] extract multiple features to characterize kin face images. The Politecnico di Torino (Polito) [26] group first combines different textural features and then selects the most relevant variables to provide an effective characterization of the samples. The LIRIS-University of Lyon (LIRIS) [26] group applies the Triangular Similarity Metric Learning method to make similar vectors parallel and dissimilar vectors opposite. Recently, a scalable similarity learning method [56] is proposed to learn a diagonal bilinear similarity model by online truncated gradient learning. Another way to reduce the appearance similarity gap is to use intermediate samples which bridge the two sides with large divergence. In [35,39,42,44], such a bridge is constructed by facial images of parents at the similar ages of their children. However, it is not easy to collect such an image set in practice.

While most of the above works focus on the bi-subject kinship verification, Ghahramani et al. [16] deal with the one-versus-multi kinship verification, i.e., predicting whether a query face image has kin relation with multiple family members, by fusing similarity of each member's facial image segments. Qin et al. [32] tackle the verification of the most basic unit that forms a family, that is tri-subject kinship verification, by incorporating the prior knowledge that a child may resemble a particular parent more than the other. Zhang et al. [53] solve the group membership prediction problem, that is predicting whether or not a collection of instances share a certain semantic property, by a probability model which consists of view-specific and view-shared random variables.

This work is different from the ones mentioned above in that we focus on the mixed bi-subject kinship verification, i.e., given a pair of query face images with genders unknown, predicting whether there is a kin relationship or not between them.

Multi-task learning (MTL) refers to the joint training of multiple problems, enforcing a common intermediate parameterization or representation. If the different problems are sufficiently related, MTL can lead to better generalization and benefit all of the tasks Evgeniou and Pontil [13,21,30,2,3,33]. To learn such common structures, some authors propose to group multiple tasks into disjoint clusters [5,12,46]. Kumar and Daume III [24] allow two tasks from different groups to overlap to some degree under the concept of latent basis tasks. Instead, Goncalves et al. [17] and Zhang and Yeung [52] propose to jointly learning the task relationship structures and the task parameters without grouping them.

In this work, we propose a new multi-task learning scheme by simultaneously learning two types of structures, one is shared amongst all the tasks and the other is unique to each task. Both are simultaneously learned in a joint framework, which enables our algorithm to utilize the common knowledge of the four tasks.

3. Mixed bi-subject kinship verification

In this section, we present our method for mixed bi-subject kinship verification. For bi-subject kinship verification problem, we mainly consider four types of kin relations, i.e., father–son (F–

S), father–daughter (F–D), mother-son (M–S) and mother–daughter (M–D). For each type of relations $t\,(1 \le t \le T,\ T = 4)$, we denote its N_t samples as $\{(x_{ti}^p, x_{ti}^c, y_{ti})\}_{i=1}^{N_t}$, where $x_{ti}^p, x_{ti}^c \in R^d$ respectively denotes the i-th sample of a parent and a child and $y_{ti} \in \{+1, -1\}$ indicates whether this child has a valid kinship relation with the corresponding parent. Our goal is to learn a function $f: (x^p, x^c) \to \{+1, -1\}$ from the training data to check whether a kinship could be established for two previously never seen images (x^p, x^c), without knowing the genders of the parent and the child.

3.1. Single task bi-subject kinship verification

For single task bi-subject kinship verification, all of the four kin relations share a common transformation matrix. Let W denote the correlation between child-parent, our learning cost function takes the following form:

$$\min_W \sum_{t=1}^{T}\sum_{i=1}^{N_t} \mathcal{L}(y_{ti}, S_W(x_{ti}^p, x_{ti}^c)) + \gamma \mathcal{R}(W) \tag{1}$$

where $\mathcal{L}(\cdot,\cdot)$ is the empirical loss function, $\mathcal{R}(W)$ is a regulariser and γ is a trade-off parameter balancing the effects of these two terms. In order to make W smooth we choose $\mathcal{R}(W) = \frac{1}{2} \| W \|_F^2$ where $\| W \|_F$ is the Frobenius norm of matrix W.

We define $\mathcal{L}(\cdot,\cdot) \equiv [a]_+ \equiv \max\{0, a\}$, $a = 1 - y_{ti} S_W(x_{ti}^p, x_{ti}^c)$. The score function $S_W(x_{ti}^p, x_{ti}^c)$ measures the similarity between the parent and the child according to their visual appearance, which takes the following bilinear function [32],

$$S_W(x^p, x^c) = (x^p)^T W x^c \equiv \langle x^p, x^c \rangle_W \tag{2}$$

where the transformation matrix W essentially captures the correlation between a parent and a child, to be learned from the training data. In what follows, we use the notation of 'STL' to denote the single task learning method for mixed bi-subject kinship verification.

3.2. From single task to multi-task

Under the viewpoint of multi-task learning, each of the four types of kin relations(F-S, F-D, M-S and M-D) is treated as a separate task but are learned together. In this work, each single task is formulated as a bilinear function f_t, i.e., $f_t(x) = (x^p)^T W_t^* x^c$, where $W_t^* \in R^{d\times d}$ is the parameter to be learnt.

Many multi-task learning methods are based on some formal definition of the notion of relatedness of the tasks. This relatedness is mostly application-dependent. For example, Hierarchical Bayes [1,4] assume that all task matrices W_t^* come from a particular probability distribution such as a Gaussian. This implies that all W_t^* are "close" to some mean function W_0. While the studies of human genetics have declared that some facial features, such as eye or hair color or dimples have been proposed to be specifically linked to paternity confidence [34]. Hair, skin and eye colors are highly heritable and visible traits in humans [38]. Thus, we think that different kinship relations may have some common genetic characteristics. Motivated by these researches, we decide to model each task in two parts,

$$W_t^* = W_0 + W_t \tag{3}$$

where the transformation matrix W_t is unique to each task, while all the tasks share a common structure W_0.

To learn these two structures, we solve the following optimization problem:

$$\min_{W_0, W_t} \mathcal{G}(W_0, W_t, \lambda, \eta) = \sum_{t=1}^{T}\sum_{i=1}^{N_t} max\{0, 1 - y_{ti}(x_{ti}^p)^T(W_0 + W_t)x_{ti}^c\} + \frac{\lambda}{2} \| W_t \|_F^2 + \frac{\eta}{2} \| W_0 \|_F^2 \tag{4}$$

where λ and η are positive regularization parameters. Intuitively, for a fixed value of η, a large value of the ratio $\upsilon = \frac{\lambda}{\eta}$ will characterize the scale of the final model. Particularly, when λ tends to infinity the problem reduces to solving one single-task learning problem as in Eq. (1). On the other hand, when η tends to infinity the problem reduces to solving the T tasks independently.

The object function $\mathcal{G}(W_0, W_t, \lambda, \eta)$ is convex with respect to W_0 and $\{W_t\}$. To obtain W_0 and $\{W_t\}$ we optimize E.q. 4 using stochastic gradient descent method, which is commonly used to minimize risk function. The gradients of W_0 and $\{W_t\}$ are calculated respectively as follows:

$$\nabla W_0 = \nabla_{W_0} \mathcal{L}(\cdot,\cdot) + \eta W_0$$

$$\nabla W_t = \nabla_{W_t} \mathcal{L}(\cdot,\cdot) + \lambda W_t \tag{5}$$

where $\nabla_{W_0} \mathcal{L}(\cdot,\cdot)$ and $\nabla_{W_t} \mathcal{L}(\cdot,\cdot)$ are the subgradients of $\mathcal{L}$ with respect to W_0 and W_t, respectively, which can be computed by:

$$\nabla_{W_0} \mathcal{L}(\cdot,\cdot) = \begin{cases} 0 & , S_{W_t^*}(x_{ti}^p, x_{ti}^c) > 1 \\ -y_{ti} x_{ti}^p x_{ti}^c & , S_{W_t^*}(x_{ti}^p, x_{ti}^c) < 1 \\ \mu_*(-y_{ti} x_{ti}^p x_{ti}^c) & , S_{W_t^*}(x_{ti}^p, x_{ti}^c) = 1 \end{cases}$$

$$\nabla_{W_t} \mathcal{L}(\cdot,\cdot) = \begin{cases} 0 & , S_{W_t^*}(x_{ti}^p, x_{ti}^c) > 1 \\ -y_{ti} x_{ti}^p x_{ti}^c & , S_{W_t^*}(x_{ti}^p, x_{ti}^c) < 1 \\ \mu_*(-y_{ti} x_{ti}^p x_{ti}^c) & , S_{W_t^*}(x_{ti}^p, x_{ti}^c) = 1 \end{cases}$$

We first optimize the objective function with respect to $\{W_t\}$ when W_0 is fixed, and then optimize it with respect to W_0 when $\{W_t\}$ is fixed. This procedure is repeated until convergence. We set $\mu = 0$ and the initial value of W_0 and $\{W_t\}$ to $\mathbf{I}$ which corresponds to the assumption that all the crossed positions of (x^p, x^c) are unrelated.

3.3. Mixed kinship verification

When the transformation matrices W_0 and $\{W_t\}$ are obtained from the training samples with genders known, it is very straightforward to make a decision for given query images with the corresponding $\{W_t\}$ for t-th task as in multi-task learning. However, in the mixed kinship verification, the genders of the query images are unknown, which means that we don't know which task to perform in the first place.

To address this, in our experiments the verification function $f(x^p, x^c)$ is modeled as the linear combination of four pieces of evidence, i.e., the confidence of (x^p, x^c) under each task, respectively

$$f(x^p, x^c) = \sum_{t=1}^{T} \beta_t S_{W_t^*}(x^p, x^c) + b \tag{6}$$

where the combination coefficients β_t are four scalars and b is the similarity threshold term. To learn these parameters, we maximize the conditional likelihood through a sigmoid function, i.e., $p(y = 1|x^p, x^c) = \sigma(f(x^p, x^c))$, with L_2 regularization added where sigmoid function σ is defined to be $\sigma(x) = \frac{1}{1+e^{-x}}$.

3.4. Multi-view multi-task learning for mixed kinship verification

Previous face verification /recognition, hand detection and palm vein recognition studies have shown that multiple feature descriptors could provide complementary information from different viewpoints [29,49–51], and hence it is desirable to utilize multiple feature information for our mixed kinship verification task. However, multiple feature descriptors generally have multiple modalities. Our MTL method is designed for single feature representation and can not be used to deal with multi features directly. A nature solution for MTL with multi-feature is to directly concatenate different features into a new feature vector and then apply MTL for verification learning. However, such an operation is not desirable and it sacrifices the diversity of different descriptors. To address this problem, we propose a new multi-view MTL (MMTL) method to seek multiple transformation matrices to perform multiple features fusion.

Assume there are V views of feature representations, and $\{(x_{ti}^{pv}, x_{ti}^{cv})\}_{i=1}^{N_t}$ denotes the v-th view set of N_t pairs of kinship images from the t-th task, where $x_{ti}^{pv} \in R^d$ and $x_{ti}^{cv} \in R^d$ are the i-th parent and child images from the v-th view, respectively, and $v = 1, 2, \ldots, V, t = 1, 2, \ldots, T$. The aim of MMTL is to learn multiple transformation matrices $W_0^1, W_0^2, \ldots, W_0^V$ and $\{W_t^1, W_t^2, \ldots W_t^V\}$ so that the property of MTL in all feature spaces can be jointly exploited.

In order to achieve this goal, we impose a nonnegative weighted vector $\beta = [\beta_1, \beta_2, \ldots, \beta_V]$ for each feature and formulate MMTL as follows:

$$\min_{W_0^v, W_t^v, \beta_v} \sum_{v=1}^{V} \beta_v \mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v)$$
$$\textit{subject to} \sum_{v=1}^{V} \beta_v = 1, \beta_v \geq 0 \tag{7}$$

where $\mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v)$ is the object function of MTL for the v-th feature, which is computed according to Eq. (4).

The trivial solution of Eq. (7) is $\beta_v = 1$, which corresponds to the minimum $\mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v)$ and $\beta_v = 0$ otherwise. This means that only the best view is selected by our method, such that the complementary information of different features has not been exploited. To address this problem, we modify β_v to be $\beta_v^u (u > 0)$, and the new objective function for MMTL is defined as:

$$\min_{W_0^v, W_t^v, \beta_v} \sum_{v=1}^{V} \beta_v^u \mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v)$$
$$\textit{subject to} \sum_{v=1}^{V} \beta_v = 1, \beta_v \geq 0 \tag{8}$$

To our best knowledge, there is no closed-form solution to Eq. (8) and there are $V_*(T+1)$ matrices to be optimized simultaneously, we solve it in an iterative manner.

First, we fix β and update W. When β is fixed, the optimization problem in Eq. (8) can be rewritten as:

$$\min_{W_0^v, W_t^v} \sum_{v=1}^{V} \beta_v^u \mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v) \tag{9}$$

We sequentially optimize $(W_0^v, \{W_t^v\})$ with the fixed $(W_0^1, \{W_t^1\})$, $(W_0^2, \{W_t^2\}), \ldots, (W_0^{v-1}, \{W_t^{v-1}\})$, $(W_0^{v+1}, \{W_t^{v+1}\}), \ldots, (W_0^V, \{W_t^V\})$. Then, Eq. (9) can be rewritten as:

$$\min_{W_0^v, W_t^v} \beta_v^u \mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v) \tag{10}$$

Now, (W_0^v, W_t^v) can be solved as that we deal with Eq. (4).

Then, we update β with the fixed W. We construct the following Lagrange function:

$$\mathcal{F}(\beta, \xi) = \sum_{v=1}^{V} \beta_v^u \mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v) - \xi\left(\sum_{v=1}^{V} \beta_v - 1\right) \tag{11}$$

Let $\frac{\partial \mathcal{F}(\beta,\xi)}{\partial \beta_v} = 0$ and $\frac{\partial \mathcal{F}(\beta,\xi)}{\partial \xi} = 0$, we have

$$u\beta_v^{u-1}\mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v) - \xi = 0 \tag{12}$$

$$\sum_{v=1}^{V} \beta_v - 1 = 0 \tag{13}$$

Combining Eqs. (12) and (13), we can obtain β_v as follows:

$$\beta_v = \frac{(1/\mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v))^{1/(u-1)}}{\sum_{v=1}^{V}(1/\mathcal{G}_v(W_0^v, W_t^v, \lambda^v, \eta^v))^{1/(u-1)}} \tag{14}$$

We repeat the above two steps until the algorithm meets a certain convergence condition. The proposed MMTL algorithm is summarized in Algorithm 1.

Algorithm 1. MMTL.

Input:

Training images: $\{(x_{ti}^{pv}, x_{ti}^{cv})\}_{i=1}^{N_t}$ be the v-th view set of N_t pairs of kinship images from the t-th task;

Parameters: regularization term λ^v, η^v, u, iteration number Nu, and convergence error τ

Output:

Symmetric transformation matrix $\{W_0^1, W_0^2, \ldots, W_0^V\}$, $\{W_t^1, W_t^2, \ldots W_t^V\}$ and wights $\beta_1, \beta_2, \ldots, \beta_V$;

1:**Step1**: **Initialization**:
2: Set $\beta_v = 1/V$, $W_0^v = \mathbf{I}$, $W_t^v = \mathbf{I}$, $t = 1, \ldots, T$, $v = 1, 2, \ldots, V$.;
3: Calculate the loss $\mathcal{L}^0(\cdot, \cdot)$;
4: **Step2**: **Local optimization**:
5: For $r = 1, 2, \ldots, Nu$, repeat
6: Compute (W_0^v, W_t^v) in Eq. (10);
7: Obtain β_v by using Eq. (14);
8: Calculate the loss $\mathcal{L}^r(\cdot, \cdot)$;
9:if $r > 2$ and $|\mathcal{L}^r - \mathcal{L}^{r-1}| < \tau$, go to Step3.;
10: **Step3**: **Output transformation matrices and weighting vector**:
11: output $\{W_0^1, W_0^2, \ldots, W_0^V\}$, $\{W_t^1, W_t^2, \ldots W_t^V\}$ and wights $\beta_1, \beta_2, \ldots, \beta_V$;.

4. Experimental results and analysis

To validate the effectiveness of the proposed approach we conducted experiments on the KinFaceW-I [27] and KinFaceW-II [27] database.

4.1. Database and experimental settings

The KinFaceW-I database [27] consists of 156 FS (Father–Son), 134 FD (Father–Daughter), 116 MS (Mother–Son) and 127 MD (Mother–Daughter) pairs, while the KinFaceW-II [27] contains 250 pairs of these bi-subject kin relations each. The major difference between KinFaceW-I and KinFaceW-II lies in that each pair of faces in KinFaceW-I comes from the same photo while from different photos in KinFaceW-II. Fig. 1 gives some examples of this database.

For each face image, we extracted three different features: (1) SIFT: We sampled SIFT descriptors on each 16×16 patch with a grid spacing of 8 pixels. Then, we concatenated these SIFT descriptors to form a 6272-dimensional feature vector. (2) LBP: We divided each image into 4×4 non-overlapping blocks, where the size of each block is 16×16. Then, we extracted a 256-dimensional uniform pattern LBP feature for each block and concatenated them to form a 4096-dimensional feature vector. (3) HOG: We divided

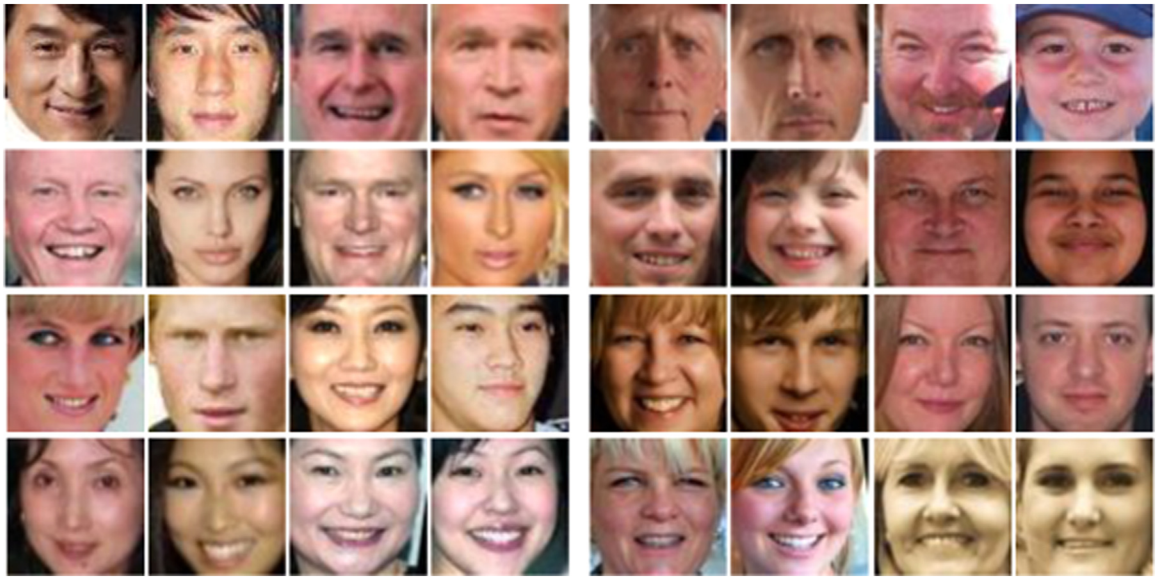


Fig. 1. Illustration of face images in the KinFaceW database. From top to bottom each row respectively shows Father–Son (F–S), Father–Daughter (F–D), Mother–Son (M–S) and Mother–Daughter (M–D) relation. The first four columns of images are from KinFaceW-I, and the last four columns of images are from KinFaceW-II.

each image into 8×8 non-overlapping blocks with the size of 2×2 pixels. Then, we extracted a 9-dimensional HOG feature for each block and concatenated them to form a 576-dimensional vector. For these three features, we apply PCA to reduce each feature into 300 dimensions to remove some noise components.

We evaluated our proposed methods with the image-restricted setting and followed the evaluation protocol as proposed in Lu et al. [26] which includes pre-specified training/testing split and independently for 5-fold cross validation. For face images in each fold, we consider all pairs of face images with the four kinship relations. Hence, the positive samples are the total true pairs of face images and the negative samples are the total false pairs of images from the four kin relations.

Note that for the parameters(λ, η, u, Nu and τ) of our method, we use 5-cross validation to set the value from a range of $\{0.001, 0.01\}$ for λ, $\{1 \times 10^{-4}$: 10: $1 \times 10^{3}\}$ or $\{1 \times 10^{-3}$: 10: $1 \times 10^{4}\}$ for η according to the value of λ, $\{1 \times 10^{-2}$: 10: $1 \times 10^{3}\}$ for u, and we set $Nu=200$, $\tau = 0.1$.

4.2. Results and analysis

4.2.1. Baseline results

We design some baselines. One is concatenating the feature vectors of two visual entities, learning a linear SVM for each kin relation, then aggregating these meta-decisions through linear SVM for the final verification judgement. We denote this method as 'concatenated+SVM'. Generally, the nonlinear SVM might get better results than the linear one, so we use a Gaussian kernel SVM for each kin relation and denote this method as 'concatenated+kernel SVM'. The third one is 'STL' as in Eq. (1).

Table 1 summarizes the results. One can see that SIFT descriptor achieves the best performance among these features. Thereafter, we will use this feature without a specific explanation.

The baseline 'Concatenated+SVM' algorithm with SIFT feature gives an accuracy of 53.3% and 55.1% on the KinFaceW-I and KinFaceW-II database, indicating that the mixed bi-subject kinship verification is a challenging problem. 'STL' method with SIFT feature improves the performance about 14% and 16% on the KinFaceW-I and KinFaceW-II database which indicates that STL method is effective. MTL method can improve the performance about 5% and 3% on the KinFaceW-I and KinFaceW-II database. The reason is

Table 1
Correct verification rates (%) for different methods on the KinFaceW-I and KinFaceW-II database.

Method	Features	KinFaceW-I	KinFaceW-II
Concatenated +SVM	LBP	51.5 ± 0.2729	52.5 ± 0.5196
	HOG	52.9 ± 0.5660	53.4 ± 0.3194
	SIFT	53.3 ± 0.2773	55.1 ± 0.6658
Concatenated +kernel SVM	LBP	52.8 ± 0.6684	53.2 ± 0.9839
	HOG	55.7 ± 0.2694	59.7 ± 0.3070
	SIFT	61.3 ± 0.5422	63.1 ± 0.4563
STL	LBP	65.7 ± 0.4343	67.7 ± 0.6319
	HOG	66.0 ± 0.3651	68.8 ± 0.4213
	SIFT	67.2 ± 0.3144	71.1 ± 0.4281
MTL (*proposed*)	LBP	69.0 ± 0.6851	70.7 ± 0.4053
	HOG	70.1 ± 0.4522	72.6 ± 0.7012
	SIFT	72.7 ± 0.2597	74.3 ± 0.3543

that MTL method utilizes the common knowledge of different kin relations as supplementary information to facilitate decision making while STL has not effectively utilized this information.

To further demonstrate the effect of our MTL method, we visualize the similarity distribution over 4000 positive and 4000 negative pairs from the KinFaceW-II database for 'STL' and MTL methods in Fig. 2. One can see that, the similarity of kinship and non-kinship pairs overlaps heavily in the learned space through 'STL' method in Fig. 2(a). However, these two similarity distribution histograms become more separate in the learned space through MTL method in Fig. 2(b), so that more discriminative information is exploited.

4.2.2. Proposed vs. different feature fusion strategies

We compare our method with three other different feature fusion strategies: (1) Single MTL(MTL): we learn a single model by using our MTL method with each single feature representation. (2) Concatenated MTL(CMTL): we first concatenate different features into a longer feature vector and then learn a single model by using our MTL method with the augmented feature representation. (3) Equal weight MTL(EMTL): we learn the model for each feature representation by our MTL method and then use the equal weight to compute the similarity of two face images.

Table 2 summarizes the results. One can see that CMTL strategy obtains the worst result among the three multi-view strategies.

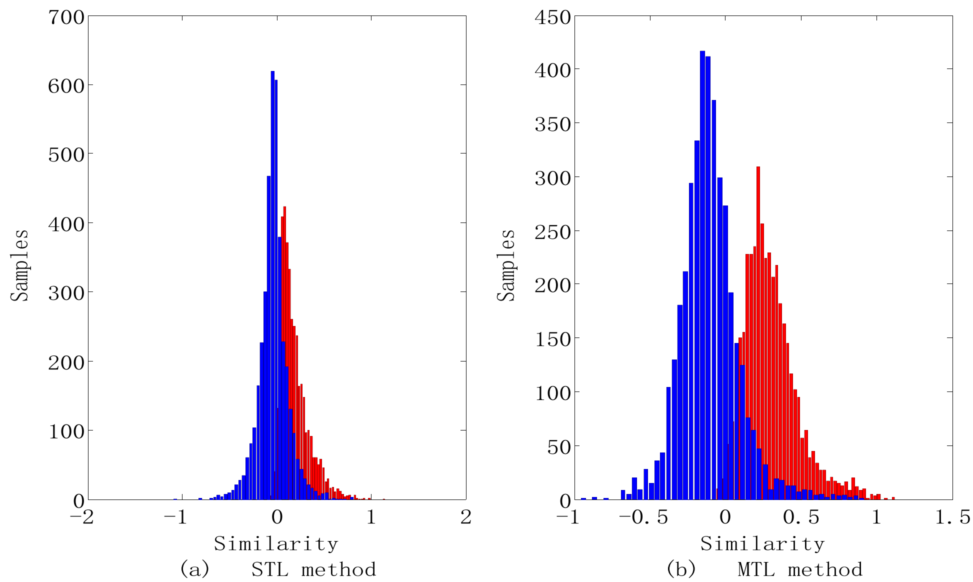


Fig. 2. Similarity distribution over 4000 positive and 4000 negative pairs from the KinFaceW-II database. (a) The distribution in the learned space through STL method. (b) The distribution in the learned space through MTL method. The histograms with the red color and the blue color represent the pairs with kinship and non-kinship, respectively. Through STL, the similarity of kinship and non-kinship pairs overlaps heavily. However, these two similarity distribution histograms become more separate in the learned space through MTL method, so that more discriminative information is exploited. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Table 2
Correct verification rates (%) for different feature fusion strategies on the KinFaceW-I and KinFaceW-II database.

Method	KinFaceW-I	KinFaceW-II
MTL(LBP)	69.0 ± 0.6851	70.7 ± 0.4053
MTL(HOG)	70.1 ± 0.4522	72.6 ± 0.7012
MTL(SIFT)	72.7 ± 0.2597	74.3 ± 0.3543
CMTL	71.6 ± 0.3446	73.0 ± 0.5143
EMTL	73.1 ± 0.3448	74.4 ± 0.5586
MMTL (*proposed*)	$\mathbf{73.7 \pm 0.3800}$	$\mathbf{77.2 \pm 0.5751}$

The reason may be that directly concatenate different features into a new longer feature vector may sacrifice the diversity of different descriptors. EMTL strategy can further improve the performance over MTL method with SIFT descriptor which indicates that different feature descriptors can provide complementary information. With the learned nonnegative weighted vector for each feature, our MMTL can obtain the best result.

4.2.3. Proposed vs. existing multi-task learning

We carry out empirical comparisons with following two competing subspace regularized multi-task learning approaches: (1) No-group MTL [25]: all tasks are assumed to be related, and the task parameters are assumed to lie in a low dimensional subspace. This is done by penalizing the Lq,1-norm of weight matrix. (2) GO-MTL [24]: a recently proposed approach that assumes each task parameter vector is a linear combination of a finite number of underlying basis tasks.

Table 3 summarizes the results. One can see that all the multi-task learning approaches are able to outperform single task learning (STL) method, indicating that common information relevant to prediction can be shared among these tasks and learning them jointly can result in better generalization performance than independently learning each task.

By incorporating over-lapping group structure, GO-MTL method improves the performance about 2.4% and 1.0% on the KinFaceW-I and KinFaceW-II database which indicates that the GO-MTL method is effective. MTL method can improve the performance about 1.1% and 1.3% on the KinFaceW-I and KinFaceW-II database. The reason is that

Table 3
Correct verification rates (%) for different methods on the KinFaceW-I and KinFaceW-II database.

Method	KinFaceW-I	KinFaceW-II
STL	67.2 ± 0.3144	71.1 ± 0.4281
No-group MTL [25]	69.2 ± 0.3747	72.0 ± 0.3683
GO-MTL [24]	71.6 ± 0.3479	73.0 ± 0.3215
MTL (*proposed*)	$\mathbf{72.7 \pm 0.2597}$	$\mathbf{74.3 \pm 0.3543}$

MTL method does not assume any grouping structure, and tasks are allowed to overlap with each other through sharing one common matrix. This is a more realistic assumption since we can have tasks in our pool that are not related enough still share some information that can be exploited for better learning.

4.2.4. Proposed vs. state-of-the-art

To the best of our knowledge, there are very few works that tackle the mixed bi-subject kinship verification problem, and it is very difficult to find an existing method directly comparable to ours.

We have compared our methods with three existing state-of-the-art separate kinship verification models, i.e., neighborhood repulsed metric learning (NRML) [27], Polito [26] and LIRIS [26].

Furthermore, considering that the similarity modeling is related to metric learning, we also include two classical metric learning algorithms, i.e., Information-theoretic metric learning (ITML) [8] and large margin nearest neighbor classification (LMNN) [41] as the base models.

Although they deal with a different problem, the image set based face verification bears some similarities to the problem of kinship verification from the respect of methodology, i.e., both involve similarity matching between multiple faces. Hence in this work, we also adopt one of the best performers on the YouTube Face database, i.e., DDML (Discriminative Deep Metric Learning) [19], to score the pairwise similarity between a child and his/her parent. Particularly, we train a deep metric learning network with three layers using our own implementation, with the threshold τ, the learning rate μ and regularization parameter λ set to be 3, 10^{-3}, 10^{-2}, respectively.

For all these methods we learn a transformation matrix for each kin relation first, then aggregate these meta-decisions through linear

Table 4
Correct verification rates (%) for different methods on the KinFaceW-I and KinFaceW-II database.

Method	KinFaceW-I	KinFaceW-II
DDML [19]	64.4 ± 0.5596	66.3 ± 0.4000
ITML [8]	64.2 ± 0.3188	67.0 ± 0.5165
LMNN [41]	67.4 ± 0.3856	67.9 ± 0.0975
NRML [27]	62.7 ± 0.3872	67.6 ± 0.4222
Polito [26]	71.1 ± 0.8273	71.8 ± 0.7365
LIRIS [26]	72.1 ± 0.2823	72.5 ± 0.5635
MTL (*proposed*)	72.7 ± 0.2597	74.3 ± 0.3543
MMTL (*proposed*)	**73.7 ± 0.3800**	**77.2 ± 0.5751**

Table 5
Correct verification rates(%) for different methods with a gender detector as a pre-processing on the KinFaceW-I and KinFaceW-II database.

Method	KinFaceW-I	KinFaceW-II
Concatenated+SVM	51.4 ± 0.5184	51.8 ± 0.3571
NRML Lu et al. [27]	61.2 ± 0.7236	67.5 ± 0.5123
Polito Lu et al. [26]	69.8 ± 0.5334	70.4 ± 0.6934
LIRIS Lu et al. [26]	71.6 ± 0.3326	71.8 ± 0.4430
DDML Hu et al. [19]	63.3 ± 0.2917	66.0 ± 0.4314
ITML Davis et al. [8]	65.9 ± 0.2815	66.0 ± 0.4393
LMNN Weinberger et al. [41]	66.6 ± 0.4209	67.0 ± 0.2031
MTL (*proposed*)	**72.7 ± 0.2597**	**74.3 ± 0.3543**

SVM for the final verification judgement. And we tune the parameters of all the compared methods to reach the best possible performance.

Table 4 summarizes the results. One can see that MTL consistently outperforms the other compared methods. This is partly due to the fact that the compared methods fail to model the relation among the two visual entities with mixed kin relationship. By contrast, 'MTL' utilizes the common knowledge of different kin relations as supplementary information to facilitate decision making and achieves better verification performance.

MMTL can improve the verification performance of MTL. The reason is MMTL can make use of multiple feature representations such that some complementary information can be utilized for our mixed kinship verification task.

4.2.5. Proposed vs. gender detector pre-processing

For mixed kinship verification, one can predict the genders of the given query images through a stand-alone gender detector and then make a decision for the image pair with the corresponding bi-subject kinship verification model. Motivated by this, we compared our MTL method with all the compared methods with a gender detector as a pre-processing with our own implementation.

Table 5 summarizes the results. It can be seen that MTL can obtain comparable or even better performance than the other compared methods. This is because we emphasize two transformation matrices – one is shared amongst all the tasks and the other is unique to each task while others have not effectively utilized this information. Another reason is that although the gender detector can achieve high performance, one must notice that only when the given query two images are correctly predicted, the pair of images can be predicted through the corresponding kin relation verification model, otherwise, the given query two images may be decided through a mismatched kin relation verification model which could not exploit this kin relation during training process.

4.2.6. Parameter analysis

We investigate the effect of the ratio $\upsilon = \frac{\lambda}{\eta}$, where we let λ fix and η vary to see the effects. Fig. 3 shows the mean verification accuracy of 'MTL' as a function of the value of υ. One can see that when υ is large we get a relatively low performance, indicating that the four kin relations should not be processed without distinction. Moreover, for KinFaceW-I the best result is obtained when $\upsilon = 1 \times 10^{-3}$ while for KinFacW-II $\upsilon = 1 \times 10^{-1}$, hence it appears that the KinFaceW-II is close to being from a single task.

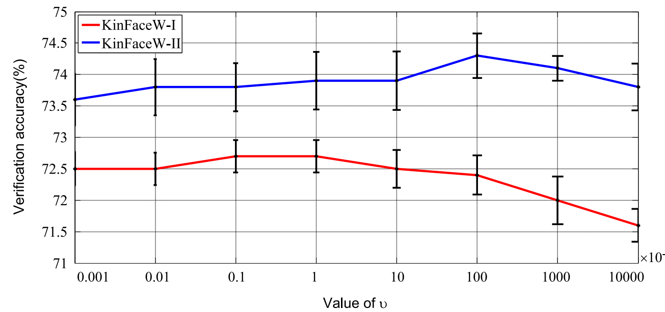


Fig. 3. Influence of the value of υ on the verification rates.

5. Conclusions

In this work, we make the first attempt to investigate the mixed bi-subject kinship verification problem. Instead of predicting whether there is a kin relation between two given entities with genders known, we try to answer this question without knowing such information. We then propose a new multi-task learning scheme to address this problem. The key idea of our method is to decompose each single task model into two parts – one is shared amongst all the tasks and the other corresponds to the task-specific structure. The model can be efficiently solved using the stochastic gradient descent method. Extensive experiments on the large scale KinFaceW kinship database show that the proposed methods significantly improve the performance of the traditional kinship verification methods. How to explore more discriminative features and combine them with our proposed MTL and MMTL methods to further improve the verification performance appears to be another interesting direction of future work.

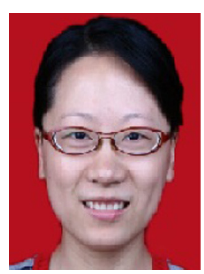
Acknowledgments

We thank the anonymous reviewers for your in-depth comments, suggestions, and corrections, which have greatly improved the manuscript. This work is partially supported by National Science Foundation of China (61373060, 61472186), Qing Lan Project and the Natural Science Foundation of the Jiangsu Higher Education Institutions of China (13KJD520002).

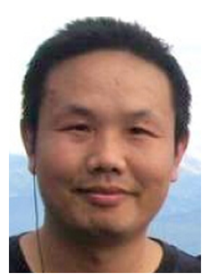
References

[1] G.M. Allenby, P.E. Rossi, Marketing models of consumer heterogeneity, J. Econom. 89 (1998) 57–78.
[2] R.K. Ando, T. Zhang, A framework for learning predictive structures from multiple tasks and unlabeled data, J. Mach. Learn. Res. 6 (2005) 1817–1853.
[3] A. Argyriou, T. Evgeniou, M. Pontil, Convex multi-task feature learning, Mach. Learn. 73 (2008) 243–272.
[4] N. Arora, G.M. Allenby, J.L. Ginter, A hierarchical bayes model of primary and secondary demand, Market. Sci. 17 (1998) 29–44.
[5] B. Bakker, T. Heskes, Task clustering and gating for bayesian multitask learning, J. Mach. Learn. Res. 4 (2003) 83–99.
[6] X. Benavent, A. Garcia-Serrano, R. Granados, J. Benavent, E. de Ves, Multimedia information retrieval based on late semantic fusion approaches: experiments on a wikipedia image collection, IEEE Trans. Multimed. 15 (2013) 2009–2021.
[7] M.F. Dal Martello, L.T Maloney, Where are kin recognition signals in the human face? J. Vis. 6 (2006) 2.
[8] J.V. Davis, B. Kulis, P. Jain, S. Sra, I.S. Dhillon, Information-theoretic metric learning, in: Proceedings of the 24th International Conference on Machine Learning, ACM, Oregon State University in Corvallis, Oregon, 2007. pp. 209–216.

[9] L.M. DeBruine, F.G. Smith, B.C. Jones, S. Craig Roberts, M. Petrie, T.D. Spector, Kin recognition signals in adult faces, Vis. Res. 49 (2009) 38–43.
[10] A. Dehghan, E.G. Ortiz, R. Villegas, M. Shah, Who do i look like? determining parent–offspring resemblance via gated autoencoders, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Columbus, Ohio, 2014.
[11] H. Dibeklioglu, A.A. Salah, T. Gevers, Like father, like son: Facial expression dynamics for kinship verification, in: 2013 IEEE International Conference on Computer Vision (ICCV), 2013, pp. 1497–1504.
[12] T. Evgeniou, C.A. Micchelli, M. Pontil, Learning multiple tasks with kernel methods, J. Mach. Learn. Res. (2005) 615–637.
[13] T. Evgeniou, M. Pontil, Regularized multi-task learning, in: Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, Seattle, WA, 2004, pp. 109–117.
[14] R. Fang, A.C. Gallagher, T. Chen, A. Loui, Kinship classification by modeling facial feature heredity, in: 2013 IEEE International Conference on Image Processing (ICIP), 2013.
[15] R. Fang, K.D. Tang, N. Snavely, T. Chen, Towards computational models of kinship verification, in: 2010 IEEE International Conference on Image Processing (ICIP), IEEE, Hong Kong, 2010, pp. 1577–1580.
[16] M. Ghahramani, W.Y. Yau, E.K. Teoh, Family verification based on similarity of individual family members facial segments, Mach. Vis. Appl. 25 (2014) 919–930.
[17] A.R. Goncalves, P. Das, S. Chatterjee, V. Sivakumar, F.J. Von Zuben, A. Banerjee, Multi-task sparse structure learning, in: Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management, ACM, New York, 2014, pp. 451–460.
[18] G. Guo, X. Wang, Kinship measurement on salient facial features, IEEE Trans. Instrum. Meas. 61 (2012) 2322–2325.
[19] J. Hu, J. Lu, Y.P. Tan, Discriminative deep metric learning for face verification in the wild, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Columbus, Ohio, 2014, pp. 1875–1882.
[20] J. Hu, J. Lu, J. Yuan, Y.P. Tan, Large margin multimetric learning for face and kinship verification in the wild, in: Proc. ACCV, 2014.
[21] A. Jalali, S. Sanghavi, C. Ruan, P.K. Ravikumar, A dirty model for multi-task learning, in: Advances in Neural Information Processing Systems, 2010, pp. 964–972.
[22] Y.G. Jiang, J. Wang, X. Xue, S.F. Chang, Query-adaptive image search with hash codes, IEEE Trans. Multimed. 15 (2013) 442–453.
[23] N. Kohli, R. Singh, M. Vatsa, Self-similarity representation of weber faces for kinship classification, in: 2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), IEEE, Arlington, Virginia, USA, 2012, pp. 245–250.
[24] A. Kumar, H. Daume III, Learning Task Grouping and Overlap in Multi-task Learning , 2012, arXiv preprint arXiv:1206.6417arXiv:1206.6417.
[25] J. Liu, S. Ji, J. Ye, SLEP: Sparse Learning with Efficient Projections. Arizona State University, 2009. http://www.public.asu.edu/jye02/Software/SLEP.
[26] J. Lu, J. Hu, V.E. Liong, X. Zhou, A. Bottino, I.U. Islam, T.F. Vieira, X. Qin, X. Tan, Y. Keller, et al. The fg 2015 kinship verification in the wild evaluation, in: 2015 IEEE International Conference on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia, 2015.
[27] J. Lu, X. Zhou, Y.P. Tan, Y. Shang, J. Zhou, Neighborhood repulsed metric learning for kinship verification, IEEE Trans. Pattern Anal. Mach. Intell. 36 (2014) 331–345.
[28] Z. Ma, Y. Yang, N. Sebe, K. Zheng, A.G. Hauptmann, Multimedia event detection using a classifier-specific intermediate representation, IEEE Trans. Multimed. 15 (2013) 1628–1637.
[29] K. Mei, L. Xu, B. Li, B. Lin, F. Wang, A real-time hand detection system based on multi-feature, Neurocomputing 158 (2015) 184–193.
[30] G. Obozinski, B. Taskar, M.I. Jordan, Joint covariate selection and joint subspace selection for multiple classification problems, Stat. Comput. 20 (2010) 231–252.
[31] S.M. Platek, D.M. Raines, G.G. Gallup Jr, F.B. Mohamed, J.W. Thomson, T. E. Myers, I.S. Panyavin, S.L. Levin, J.A. Davis, L. Fonteyn, et al., Reactions to children's faces: males are more affected by resemblance than females are, and so are their brains, Evolut. Hum. Behav. 25 (2004) 394–405.
[32] X. Qin, X. Tan, S. Chen, Tri-subject kinship verification: understanding the core of a family, 2015, arXiv preprint arXiv:1501.02555arXiv:1501.02555.
[33] P. Rai, H. Daume, Infinite predictor subspace models for multitask learning, in: International Conference on Artificial Intelligence and Statistics, 2010, pp. 613–620.
[34] F. Salter, Carrier females and sender males: an evolutionary hypothesis linking female attractiveness, family resemblance, and paternity confidence, Ethol. Sociobiol. 17 (1996) 211–220.
[35] M. Shao, D. Kit, Y. Fu, Generalized transfer subspace learning through low-rank constraint, Int. J. Comput. Vis. 109 (2014) 74–93.
[36] M. Shao, S. Xia, Y. Fu, Identity and kinship relations in group pictures, in: Human-Centered Social Media Analytics, Springer, New York, 2014, pp. 175–190.
[37] E. Spyromitros-Xioufis, S. Papadopoulos, I. Kompatsiaris, G. Tsoumakas, I. Vlahavas, A comprehensive study over vlad and product quantization large-scale image retrieval, IEEE Trans. Multimed. 16 (2014) 1713–1728.
[38] P. Sulem, D.F. Gudbjartsson, S.N. Stacey, A. Helgason, T. Rafnar, K.P. Magnusson, A. Manolescu, A. Karason, A. Palsson, G. Thorleifsson, et al., Genetic determinants of hair, eye and skin pigmentation in Europeans, Nat. Genet. 39 (2007) 1443–1452.
[39] N. Syed, B.K.P.P.Q. Shareq, et al., Understanding familial relationship in an image, Int. J. Scient. Res. Educ. (2014) 2.
[40] G. Wang, A. Gallagher, J. Luo, D. Forsyth, Seeing people in social context: Recognizing people and social relationships, in: Computer Vision–ECCV 2010, Springer, New York, 2010, pp. 169–182.
[41] K. Weinberger, J. Blitzer, L. Saul, Distance metric learning for large margin nearest neighbor classification, Adv. Neural Inf. Process. Syst. 18 (2006) 1473.
[42] S. Xia, M. Shao, Y. Fu, Kinship verification through transfer learning, in: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, Volume Three, AAAI Press, San Francisco, 2011, pp. 2539–2544.
[43] Xia, S., Shao, M., Fu, Y., 2012a. Toward kinship verification using visual attributes, in: 2012 International Conference on Pattern Recognition (ICPR), IEEE, Tsukuba Science city, Japan, pp. 549–552.
[44] S. Xia, M. Shao, J. Luo, Y. Fu, Understanding kin relationships in a photo, IEEE Trans. Multimed. 14 (2012) 1046–1056.
[45] Z. Xu, Y. Zhang, L. Cao, Social image analysis from a non-iid perspective, IEEE Trans. Multimed. 16 (2014) 1986–1998.
[46] Y. Xue, X. Liao, L. Carin, B. Krishnapuram, Multi-task learning for classification with Dirichlet process priors, J. Mach. Learn. Res. 8 (2007) 35–63.
[47] H. Yan, J. Lu, W. Deng, X. Zhou, Discriminative multi-metric learning for kinship verification, IEEE Trans. Inf. Forens. Secur. 9 (2014) 1169–1178.
[48] H. Yan, J. Lu, X. Zhou, Prototype-based discriminative feature learning for kinship verification, IEEE Trans. Cybern. (2014) 1.
[49] H. Yan, J. Lu, X. Zhou, Y. Shang, Multi-feature multi-manifold learning for single-sample face recognition, Neurocomputing 143 (2014) 134–143.
[50] X. Yan, W. Kang, Q. Wu, F. Deng, Palm vein recognition based on multi-sampling and feature-level fusion, Neurocomputing 151 (2015) 798–807.
[51] J. Yang, J.y. Yang, D. Zhang, J.f. Lu, Feature fusion: parallel strategy vs. serial strategy, Pattern Recognit. 36 (2003) 1369–1381.
[52] Y. Zhang, D.Y. Yeung, A convex formulation for learning task relationships in multi-task learning. arXiv preprint arXiv:1203.3536arXiv:1203.3536, 2012.
[53] Z. Zhang, Y. Chen, V. Saligrama, Group membership prediction. arXiv preprintarXiv:1509.04783arXiv:1509.04783, 2015.
[54] X. Zhou, J. Hu, J. Lu, Y. Shang, Y. Guan, Kinship verification from facial images under uncontrolled conditions, in: Proceedings of the 19th ACM International Conference on Multimedia, ACM, 2011, pp. 953–956.
[55] X. Zhou, J. Lu, J. Hu, Y. Shang, Gabor-based gradient orientation pyramid for kinship verification under uncontrolled environments, in: Proceedings of the 20th ACM International Conference on Multimedia, ACM, 2012, pp. 725–728.
[56] X. Zhou, H. Yan, Y. Shang, Kinship verification from facial images by scalable similarity fusion, Neurocomputing (2016).
[57] O. Zoidi, A. Tefas, N. Nikolaidis, I. Pitas, Person identity label propagation in stereo videos, IEEE Trans. Multimed. 16 (2014) 1358–1368.



Xiaoqian Qin is a Ph.D. candidate at Nanjing University of Aeronautics and Astronautics. She is a Lecturer at Huaiyin Normal University. Her research interests include computer vision and machine learning.



Xiaoyang Tan is a Professor in the Department of Computer Science and Technology at the Nanjing University of Aeronautics and Astronautics. He obtained his Ph.D. from Nanjing University in 2005. From September 2006 to October 2007, he worked as a Postdoctoral Researcher in the LEAR team at INRIAR Rhone-Alpes in Grenoble, France. His research interests include face recognition, computer vision, machine learning, and artificial intelligence. He has published over 40 conference and journal papers in these areas.

Songcan Chen received the B.S. degree in mathematics from Hangzhou University (now merged into Zhejiang University), Zhejiang, China, the M.S. degree in computer applications from Shanghai Jiaotong University, Shanghai, China, and the Ph.D. degree in communication and information systems from the Nanjing University of Aeronautics and Astronautics (NUAA), Nanjing, China, in 1983, 1985, and 1997, respectively. He was at NUAA in January 1986. Since 1998, he has been a full-time Professor with the Department of Computer Science and Technology, NUAA. He has authored or co-authored over 200 scientific peer-reviewed papers. His current research interests include pattern recognition, machine learning and neural computing.