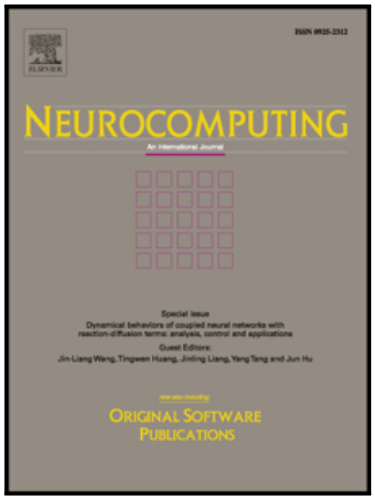
Journal Pre-proof



Tensor Cross-view Quadratic Discriminant Analysis for Kinship Verification in the Wild

Oualid Laiadi, Abdelmalik Ouamane, Abdelhamid Benakcha, Abdelmalik Taleb-Ahmed, Abdenour Hadid

PII: S0925-2312(19)31435-3
DOI: https://doi.org/10.1016/j.neucom.2019.10.055
Reference: NEUCOM 21402

To appear in: *Neurocomputing*

Received date: 10 December 2018
Revised date: 3 August 2019
Accepted date: 5 October 2019

Please cite this article as: Oualid Laiadi, Abdelmalik Ouamane, Abdelhamid Benakcha, Abdelmalik Taleb-Ahmed, Abdenour Hadid, Tensor Cross-view Quadratic Discriminant Analysis for Kinship Verification in the Wild, *Neurocomputing* (2019), doi: https://doi.org/10.1016/j.neucom.2019.10.055

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier B.V.

Journal Pre-proof

Tensor Cross-view Quadratic Discriminant Analysis for Kinship Verification in the Wild

Oualid Laiadi[a,d], Abdelmalik Ouamane[b], Abdelhamid Benakcha[c], Abdelmalik Taleb-Ahmed[d], Abdenour Hadid[e]

[a]*Laboratory of LESIA, University of Biskra, Algeria*
[b]*University of Biskra, Algeria*
[c]*Laboratory of LGEB, University of Biskra, Algeria*
[d]*IEMN DOAE UMR CNRS 8520 Laboratory, Polytechnic University of Hauts-de-France, France*
[e]*Center for Machine Vision and Signal Analysis, University of Oulu, Finland*

Abstract

This paper presents a new Tensor Cross-view Quadratic Discriminant Analysis (TXQDA) method based on the XQDA method for kinship verification in the wild. Many researchers used metric learning methods and have achieved reasonably good performance in kinship verification, none of these methods looks at the kinship verification as a cross-view matching problem. To tackle this issue, we propose a tensor cross-view method to train multilinear data using local histograms of local features descriptors. Therefore, we learn a hierarchical tensor transformation to project each pair face images into the same implicit feature space, in which the distance of each positive pair is minimized and that of each negative pair is maximized. Moreover, TXQDA was proposed to separate the multifactor structure of face images (i.e. kinship, age, gender, expression, illumination and pose) from different dimensions of the tensor. Thus, our TXQDA achieves better classification results through discovering a lowdimensional tensor subspace that enlarges the margin of different kin relation classes. Experimental evaluation on five challenging databases namely Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases, show that the proposed TXQDA significantly outperforms the current state of the art.

*Corresponding author

Journal Pre-proof

Journal Pre-proof

Keywords: kinship verification, cross-view, Tensor XQDA.

1. Introduction

Kinship verification from face images which is increasingly attracting the attention of the research community is an emerging research topic in computer vision. Checking if two persons are from the same family or not can be automatically verified through facial images. Learning and extracting the face similarity between family members is challenging. Many potential applications, such as creation of family trees, family album organization, image annotation, finding missing children and forensics, are targeted by kinship verification. Although a DNA test is the most reliable way for kinship verification, it cannot be used in many situations. Kinship verification via faces can typically be done in video surveillance scenes. In addition to the difficulties generally encountered in face verification in the wild (i.e. facial images captured under uncontrolled environments without any restrictions in terms of pose, lighting, background, expression, and partial occlusion), kinship verification adds another layer of difficulty which is far from being easy. Kinship verification deals with facial images which belong inevitably to different persons with a significant age difference and in some cases from a different gender. Further, the face attributes of persons of the same family may show a large dissimilarity whereas pair faces of persons with no kinship may look similar. All these challenges increase the difficulties of the kinship verification problem.

Many researchers [1, 2, 3, 4, 5, 6, 7, 8, 9, 10] used metric learning methods and have achieved reasonably good performance in kinship verification, but none of these methods tackle the kinship verification as a cross-view matching problem.

To our best knowledge, our work is the first effort that tackles the kinship verification problem with a method used in the cross-view matching problem that arise from many applications like heterogeneous face recognition [11] and viewpoint invariant person re-identification [12]. The Cross-view Quadratic

Journal Pre-proof

Discriminant Analysis (XQDA) [13] method shows the best performances in person re-identification field. Motivated by this research, we propose Tensor Cross-view Quadratic Discriminant Analysis (TXQDA) to analyze the multifactor structure of face images which is related to kinship, age, gender, expression, illumination and pose.

In our framework, the set of face images are represented as a third-order tensor based on local histogram features of the local descriptors, Multi-Scale Local Phase Quantization [14], and Multi-Scale local Binarised Statistical Image Features [15]. The contributions of this work are summarized as follows:

1. We tackle for the first time the kinship verification problem as a cross-view matching problem because every kin relation is typically viewpoint changes from two face images belonging to two different persons.

2. We propose a robust automated facial verification framework suitable for kinship verification, from face images captured in unconstrained environments. The face data is represented as a high order tensor based on the combination of different local features in order to provide a more powerful face model.

3. We propose a novel method for dimensionality reduction and classification, called Tensor Cross-view Quadratic Discriminant Analysis (TXQDA), which preserves the data structure, enlarges the margin between samples, helps lighten the small sample size problem and reduced the computational cost.

4. We extensively evaluate our TXQDA method against the state-of-the-art methods using five challenging kinship databases namely Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW.

The paper is organized as follows: Section 2 presents some related works on multilinear subspace methods and features extraction. Section 3 describes the proposed Tensor Cross-view Quadratic Discriminant Analysis. Our tensor kinship verification pipeline is presented in Section 4. The experiments and

Journal Pre-proof

results are given in Section 5. Finally, concluding remarks are given in Section 6.

2. Related work

Over the past decade, many algorithms have been proposed to improve kinship verification from faces. Metric learning methods which are considered as one of the greatest applied dimensionality reduction methods are widely used for face recognition. The objective of these methods is to learn and reduce the high dimensional feature subspace into a lower and discriminated subspace which leads to a better separation between the classes [16, 17]. Among the most important approaches, are the linear dimensionality reduction methods that includes the Principal Component Analysis (PCA) [18] and the Linear Discriminant Analysis (LDA) [19]. PCA's aim is to increase the transformed features variance, which extracted in the projected subspace. However, LDA increases the inter-class covariance while decreasing the intra-class covariance in the projected subspace. These traditional algorithms deal with the transformed input face image, from matrix of size x $\times$ y to get high-dimensional features vector 1D by assembling the columns or rows of the images [20]. Inevitably, by this transformation to 1D-vector some important discriminative structural information can be lost. Later on, some researches [21, 22, 23] proved that high order data representation in multilinear subspace methods based on tensor analysis, lead to excellent performances in the face recognition. The linear dimensionality reduction methods have been extended into multilinear subspace methods which are based on the high order tensor representation replaced by their vectorized forms [20, 24, 25, 26, 27, 21, 22, 28]. In another hand, multilinear transformations methods have the number of indices (N) that defines its order, process the multifactor structure of the face images. In this way, the multiple factors (i.e. expression, illumination and pose) can be separated from different dimensions of the tensor [29].

The multilinear principal component analysis (MPCA) [30] proposed as a

Journal Pre-proof

multilinear extension of the PCA, defines a multilinear projection that projects the original structural tensor into a lower dimensional tensor subspace while conserving the disparity in the original samples and add more intrinsic structural information [22]. Moreover, LDA has been extended into many multilinear variants, such as Multilinear Discriminant Analysis (MDA) [23] which has a multiple interrelated subspaces that can collaborate to discriminate different classes. Also, MDA is extended to different approaches such as General Tensor Discriminant Analysis (GTDA) [31], Uncorrelated Multilinear Discriminant Analysis (UMDA) [22], Three-dimensional Modular Discriminant Analysis (3DMDA) [27] and Sparse Tensor Discriminant Analysis (STDA) [20].

Usually, geometrical face features like nose, chin structure, mouth and eyes are affected by the human facial appearance variations (facial expressions, poses variations, conditions of illumination). The research in face recognition topic has more and more turned to use descriptor-based approaches due to their robustness to the aforementioned facial appearance variations [32]. Recently, histograms-based face representations have been widely used to improve feature extraction [33]. Arguably, LPQ (Local Phase Quantization) [14] and BSIF (Binarized Statistical Image Features) [15] which have been used in this study, are one of the most prominent ones among these new local descriptors. LPQ descriptor is constructed to keep only the local information that is invariant to blur. It is designed by Ojansivu *et al.* [14] in 2008. LPQ extracts local phase by quantizing the Fourier transform phase in M by M neighborhood window at each pixel of the input image. Evidently, the Fourier transform is calculated for only four frequencies. Thereafter, a simple scalar quantizer is used to extract information from each phase coefficients by observing the signs of the real and imaginary parts. The number of bits per point is eight coefficients, which are represented as integer value between 0 and 255 using a simple binary coding. One year later, the authors in [34] proposed a Multi-Scale LPQ (MSLPQ) features extraction to improve the discriminative power of LPQ. Different neighborhood window size is applied to the input image and the LPQ labeled images are combined by concatenation or fusion to represent the input image. Inspired

Journal Pre-proof

by LPQ methodology, Kannala and Rahtu [15] proposed the Binarized Statistical Image Features (BSIF) in 2012. BSIF is a learning based-descriptor that assigns to each pixel in the input image a binary code using its response to a set of linear filters that are automatically learned using the statistical properties of 13 natural images [35]. Independent Component Analysis (ICA) is used for learning the linear filters by maximizing the statistical independence between the responses of each filter and given the image patches [15].

3. Tensor Cross-view Quadratic Discriminant Analysis

The variables and mathematical notations that we used in our work are as follows : Lowercase and uppercase symbols (e.g., i, j, F, N and V) indicate scalars; Bold lowercase symbols (e.g., $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{z}$) indicate vectors; italic uppercase symbols (e.g., U, X, Y and W) indicate matrices; bold italic uppercase symbols (e.g., $\boldsymbol{X}$, $\boldsymbol{Y}$, and $\boldsymbol{Z}$) indicate tensors. A tensor is explained as a multidimensional array [20, 23]. N is considered the order of the tensor and $\boldsymbol{X}$ is called an N^{th}-order tensor. $\mathrm{I_k}$, $1 \leq \mathrm{k} \leq \mathrm{N}$, is the dimension of the k^{th} mode.

3.1. Cross-view Quadratic Discriminant Analysis (XQDA)

XQDA [13] is the extended method of the Bayesian face [36] and KISSME [37] approaches to cross-view metric learning, where considered to learn a subspace $W = (\mathrm{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_\mathrm{r}) \in \Re^{\mathrm{d}\times\mathrm{r}}$ with cross-view (i.e. Parent-Child) data, and learn a distance function in the r dimensional subspace for the cross-view similarity measure at the same time. We assume that we have a cross-view training set $\{X, Z\}$ of c classes, in which $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_\mathrm{n}) \in \Re^{\mathrm{d}\times\mathrm{n}}$ includes n samples in a d-dimensional space from one view (i.e. Parents samples), $Z = (\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_\mathrm{m}) \in \Re^{\mathrm{d}\times\mathrm{m}}$ includes m samples in the same d-dimensional space but from the other view (i.e. Children samples). Note that Z is the same with X in the single-view matching scenario. Considering a subspace W, the distance function in the r dimensional subspace is computed as:

$$\mathrm{d}_W(\mathrm{x}, \mathrm{z}) = (\mathrm{x} - \mathrm{z})^T W(\Sigma_I^{'-1} - \Sigma_E^{'-1})W^T(\mathrm{x} - \mathrm{z}) \tag{1}$$

Journal Pre-proof

Where $\Sigma_I^{'} = W^T \Sigma_I W$ and $\Sigma_E^{'} = W^T \Sigma_E W$. Then, we learn a kernel matrix $M(W) = W(\Sigma_I^{'-1} - \Sigma_E^{'-1})W^T$. In [13] the projection direction W is optimized such that $\Sigma_E^{'}/\Sigma_I^{'}$ is maximized. Consequently, $\Sigma_E^{'}/\Sigma_I^{'}$ corresponds to the Generalized Rayleigh Quotient:

$$J(W) = \frac{W^T \Sigma_E W}{W^T \Sigma_I W} \tag{2}$$

The two covariance matrices Σ_E and Σ_I are computed as follow:

$$\mathrm{n_I}\Sigma_I = \tilde{X}\tilde{X}^T + \tilde{Z}\tilde{Z}^T - SR^T - RS^T \tag{3}$$

Where $\tilde{X} = (\sqrt{\mathrm{m_1}}\mathbf{x}_1, \sqrt{\mathrm{m_1}}\mathbf{x}_2, \ldots, \sqrt{\mathrm{m_1}}\mathbf{x}_{\mathrm{n_1}}, \ldots, \sqrt{\mathrm{m_c}}\mathbf{x}_{\mathrm{n}})$,
$\tilde{Z} = (\sqrt{\mathrm{n_1}}\mathbf{z}_1, \sqrt{\mathrm{n_1}}\mathbf{z}_2, \ldots, \sqrt{\mathrm{n_1}}\mathbf{z}_{\mathrm{m_1}}, \ldots, \sqrt{\mathrm{n_c}}\mathbf{z}_{\mathrm{m}})$,
$S = (\sum_{\mathrm{y_i}=1} \mathbf{x}_{\mathrm{i}}, \sum_{\mathrm{y_i}=2} \mathbf{x}_{\mathrm{i}}, \ldots, \sum_{\mathrm{y_i}=\mathrm{c}} \mathbf{x}_{\mathrm{i}})$, $R = (\sum_{\mathrm{l_j}=1} \mathbf{z}_{\mathrm{j}}, \sum_{\mathrm{l_j}=2} \mathbf{z}_{\mathrm{j}}, \ldots, \sum_{\mathrm{l_j}=\mathrm{c}} \mathbf{z}_{\mathrm{j}})$, $\mathrm{y_i}$ and $\mathrm{l_j}$ are class labels, $\mathrm{n_i}$ is the number of samples in class i of X, and $\mathrm{m_i}$ is the number of samples in class i of Z.

$$\mathrm{n_E}\Sigma_E = \mathrm{m}XX^T + \mathrm{n}ZZ^T - \mathbf{s}\mathbf{r}^T - \mathbf{r}\mathbf{s}^T - \mathrm{n_I}\Sigma_I \tag{4}$$

Where $\mathbf{s} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathbf{x}_{\mathrm{i}}$ and $\mathbf{r} = \sum_{\mathrm{j}=1}^{\mathrm{m}} \mathbf{z}_{\mathrm{j}}$

3.2. Tensor Cross-view Quadratic Discriminant Analysis (TXQDA)

Let a Tensor cross-view training set $\{\boldsymbol{X}, \boldsymbol{Z}\}$ of c classes, where: $\boldsymbol{X} \in \Re^{\mathrm{I_1} \times \mathrm{I_2} \times \cdots \times \mathrm{I_N} \times \mathrm{n}}$ contains n samples of one view (Parents samples) and $\boldsymbol{Z} \in \Re^{\mathrm{I_1} \times \mathrm{I_2} \times \cdots \times \mathrm{I_N} \times \mathrm{m}}$ contains m samples of other view (Children samples). The goal of our TXQDA is the calculation of N projection matrices ($W_1 \in \Re^{\mathrm{I_1} \times \mathrm{I_1^{'}}}, W_2 \in \Re^{\mathrm{I_2} \times \mathrm{I_2^{'}}}, \ldots, W_{\mathrm{N}} \in \Re^{\mathrm{I_N} \times \mathrm{I_N^{'}}}$). Thus, we calculate one projection matrix for each tensor mode. The objective function of XQDA 2 is transformed into:

$$J(W_{\mathrm{k}}) = \frac{W_{\mathrm{k}}^T \Sigma_E^{\mathrm{k}} W_{\mathrm{k}}}{W_{\mathrm{k}}^T \Sigma_I^{\mathrm{k}} W_{\mathrm{k}}} \tag{5}$$

Journal Pre-proof

We calculate the two covariance matrices Σ_E^{k} and Σ_I^{k} for each k mode by:

$$\mathrm{n_I}\Sigma_I = \sum_{\mathrm{p}=1}^{\prod_{\mathrm{o}\neq\mathrm{k}}\mathrm{I_o}} \mathrm{n_I}\Sigma_I^{\mathrm{p}}, \mathrm{n_I}\Sigma_I^{\mathrm{p}} = \tilde{X}^{\mathrm{k,p}}(\tilde{X}^{\mathrm{k,p}})^T + \tilde{Z}^{\mathrm{k,p}}(\tilde{Z}^{\mathrm{k,p}})^T - S^{\mathrm{k,p}}(R^{\mathrm{k,p}})^T - R^{\mathrm{k,p}}(S^{\mathrm{k,p}})^T \tag{6}$$

Where $\tilde{X}^{\mathrm{k,p}} = (\sqrt{\mathrm{m_1}}\mathbf{x}_1^{\mathrm{k,p}}, \sqrt{\mathrm{m_1}}\mathbf{x}_2^{\mathrm{k,p}}, \ldots, \sqrt{\mathrm{m_1}}\mathbf{x}_{\mathrm{n_1}}^{\mathrm{k,p}}, \ldots, \sqrt{\mathrm{m_c}}\mathbf{x}_{\mathrm{n}}^{\mathrm{k,p}})$,
$\tilde{Z}^{\mathrm{k,p}} = (\sqrt{\mathrm{n_1}}\mathbf{z}_1^{\mathrm{k,p}}, \sqrt{\mathrm{n_1}}\mathbf{z}_2^{\mathrm{k,p}}, \ldots, \sqrt{\mathrm{n_1}}\mathbf{z}_{\mathrm{m_1}}^{\mathrm{k,p}}, \ldots, \sqrt{\mathrm{n_c}}\mathbf{z}_{\mathrm{m}}^{\mathrm{k,p}})$,
$S^{\mathrm{k,p}} = (\sum_{\mathrm{y_i}=1} \mathbf{x}_{\mathrm{i}}^{\mathrm{k,p}}, \sum_{\mathrm{y_i}=2} \mathbf{x}_{\mathrm{i}}^{\mathrm{k,p}}, \ldots, \sum_{\mathrm{y_i}=\mathrm{c}} \mathbf{x}_{\mathrm{i}}^{\mathrm{k,p}})$, $R^{\mathrm{k,p}} = (\sum_{\mathrm{l_j}=1} \mathbf{z}_{\mathrm{j}}^{\mathrm{k,p}}, \sum_{\mathrm{l_j}=2} \mathbf{z}_{\mathrm{j}}^{\mathrm{k,p}}, \ldots, \sum_{\mathrm{l_j}=\mathrm{c}} \mathbf{z}_{\mathrm{j}}^{\mathrm{k,p}})$,

Where, for all presentations, $\mathbf{x}^{\mathrm{k,p}}$ and $\mathbf{z}^{\mathrm{k,p}}$ are the p^{th} column vectors of the k-mode unfolded matrices X^{k} and Z^{k} of sample tensors $\mathcal{X}$ and $\boldsymbol{\mathcal{Z}}$, respectively.

$$\mathrm{n_E}\Sigma_E = \sum_{\mathrm{p}=1}^{\prod_{\mathrm{o}\neq\mathrm{k}}\mathrm{I_o}} \mathrm{n_E}\Sigma_E^{\mathrm{p}}, \mathrm{n_E}\Sigma_E^{\mathrm{p}} = \mathrm{m}X^{\mathrm{k,p}}(X^{\mathrm{k,p}})^T + \mathrm{n}Z^{\mathrm{k,p}}(Z^{\mathrm{k,p}})^T - \mathbf{s}^{\mathrm{k,p}}(\mathbf{r}^{\mathrm{k,p}})^T - \mathbf{r}^{\mathrm{k,p}}(\mathbf{s}^{\mathrm{k,p}})^T - \mathrm{n_I}\Sigma_I^{\mathrm{p}} \tag{7}$$

Where $X^{\mathrm{k,p}} = (\mathbf{x}_1^{\mathrm{k,p}}, \mathbf{x}_2^{\mathrm{k,p}}, \ldots, \mathbf{x}_{\mathrm{n_1}}^{\mathrm{k,p}}, \ldots, \mathbf{x}_{\mathrm{n}}^{\mathrm{k,p}})$, $Z^{\mathrm{k,p}} = (\mathbf{z}_1^{\mathrm{k,p}}, \mathbf{z}_2^{\mathrm{k,p}}, \ldots, \mathbf{z}_{\mathrm{m_1}}^{\mathrm{k,p}}, \ldots, \mathbf{z}_{\mathrm{m}}^{\mathrm{k,p}})$,
$\mathbf{s}^{\mathrm{k,p}} = \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathbf{x}_{\mathrm{i}}^{\mathrm{k,p}}$ and $\mathbf{r}^{\mathrm{k,p}} = \sum_{\mathrm{j}=1}^{\mathrm{m}} \mathbf{z}_{\mathrm{j}}^{\mathrm{k,p}}$

Now that the solution for one mode is known, the optimization problem in equation 5 can be solved iteratively. The projection matrices $W_1, W_2, \ldots, W_{\mathrm{N}}$ are first initialized to identity. At each iteration $W_1, W_2, \ldots, W_{\mathrm{k}-1}, W_{\mathrm{k}+1}, \ldots W_{\mathrm{N}}$ are hypothetical known and W_{k} is estimated. Set: $\boldsymbol{\mathcal{U}} = \boldsymbol{\mathcal{X}} \times_1 W_1 \ldots \times_{\mathrm{k}-1} W_{\mathrm{k}-1} \times_{\mathrm{k}+1} W_{\mathrm{k}+1} \ldots \times_{\mathrm{N}} W_{\mathrm{N}}$ and $\boldsymbol{\mathcal{Y}} = \boldsymbol{\mathcal{Z}} \times_1 W_1 \ldots \times_{\mathrm{k}-1} W_{\mathrm{k}-1} \times_{\mathrm{k}+1} W_{\mathrm{k}+1} \ldots \times_{\mathrm{N}} W_{\mathrm{N}}$ are replaced in equation 5 by $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Z}}$. The new equation can be solved by the generalized eigenvalue decomposition problem:

$$\Sigma_E^{\mathrm{k}} W_{\mathrm{k}} = \Lambda_{\mathrm{k}} \Sigma_I^{\mathrm{k}} W_{\mathrm{k}} \tag{8}$$

Where, W_{k} is the eigenvectors matrix and Λ_{k} the eigenvalues matrix.

The iterative process of TXQDA breaks up on the recognition of one of the following situations: i) The number of iterations reaches a predefined maximum; or ii) the difference of the estimated projection between two consecutive iterations is less than a threshold, $\left\|W_{\mathrm{k}}^{\mathrm{iter}} - W_{\mathrm{k}}^{\mathrm{iter}-1}\right\| < \mathrm{I_k}\mathrm{I_k}\epsilon$, where $\mathrm{I_k}$ is the k

Journal Pre-proof

mode dimension of $W_{\mathrm{k}}^{\mathrm{iter}}$. The number of iterations, for our TXQDA algorithm, is empirically tuned and the better value is $\mathrm{Iteration}_{\max} = 2$.

We summarize the advantages of our algorithm, Tensor cross-view quadratic discriminant analysis (TXQDA), as follows:

1. TXQDA preserves the data structure, where these data stacked in a tensor mode providing the maximum extraction of information. Unlike in the case of XQDA method, the feature vectors are purely concatenated neglecting the natural structure of data

2. TXQDA also helps lightening the small sample size problem. This is an intrinsic limitation of the XQDA when applying the histograms concatenation of local descriptors for all face blocks (the features length is larger than the number of training samples)

3. TXQDA is a cross-view dimensionality reduction method. It can obviate the curse of dimensionality dilemma by using higher order tensors and k-mode optimization approach, where the latter is performed in a much lower-dimension feature space than the traditional vector-based methods, such as XQDA, do.

4. Many more feature dimensions are available in TXQDA than in XQDA because the available feature dimension of XQDA is theoretically limited by the number of classes in the data, whereas the TXQDA is not.

5. TXQDA reduces the computational cost to a large extent, as the k-mode optimization in each step is performed on a feature space of smaller size.

Consequently, the classification with the proposed TXQDA is better than XQDA. The entire procedure for the proposed Tensor Cross-view Quadratic Discriminant Analysis (TXQDA) is provided in Algorithm 1. The input of this algorithm is defined as follow:

- The tensor $\boldsymbol{X} \in \Re^{\mathrm{I}_1 \times \mathrm{I}_2 \times \cdots \times \mathrm{I}_{\mathrm{N}} \times \mathrm{n}}$ contains n samples of one view (Parents samples).

Journal Pre-proof

- The tensor $\boldsymbol{Z} \in \Re^{\mathrm{I}_1 \times \mathrm{I}_2 \times \cdots \times \mathrm{I}_\mathrm{N} \times \mathrm{m}}$ contains m samples of other view (Children samples).
- $\mathrm{Iteration}_{\max}$ is the maximal number of iterations.
- The final lower dimensions: $\acute{\mathrm{I}_1} \times \acute{\mathrm{I}_2} \times \cdots \times \acute{\mathrm{I}_\mathrm{N}}$.

Whereas the output can defined as follow:

- The projection matrices $W_\mathrm{k} = W_\mathrm{k}^{\mathrm{iter}} \in \Re^{\mathrm{I_k} \times \acute{\mathrm{I}_\mathrm{k}}}, \mathrm{k} = 1, \cdots, \mathrm{N}$

4. Proposed Tensor Kinship verification pipeline

In this section, we explain the details of employing the proposed TXQDA for kinship verification from pairs of face images. As depicted in Fig. 1, the block diagram of the proposed approach consists of three essential components: feature extraction, tensor subspace transformation and comparison. We focus in this work on subspace transformation and the feature extraction based multiple scales local descriptor.

Journal Pre-proof

Journal Pre-proof

Algorithm 1 Tensor Cross-view Quadratic Discriminant Analysis (TXQDA)

Input:

- $\boldsymbol{X} \in \Re^{\mathrm{I}_1 \times \mathrm{I}_2 \times \cdots \times \mathrm{I}_\mathrm{N} \times \mathrm{n}}$
- $\boldsymbol{Z} \in \Re^{\mathrm{I}_1 \times \mathrm{I}_2 \times \cdots \times \mathrm{I}_\mathrm{N} \times \mathrm{m}}$
- $\mathrm{Iteration}_{\max}$
- $\acute{\mathrm{I}}_1 \times \acute{\mathrm{I}}_2 \times \cdots \times \acute{\mathrm{I}}_\mathrm{N}$

Output:

- $W_\mathrm{k} = W_\mathrm{k}^{\mathrm{iter}} \in \Re^{\mathrm{I}_\mathrm{k} \times \acute{\mathrm{I}}_\mathrm{k}}, \mathrm{k} = 1, \cdots, \mathrm{N}$

Algorithm:

1. **Initialization:** $W_1^0 = I_{\mathrm{I}_1}, W_2^0 = I_{\mathrm{I}_2}, \cdots, W_\mathrm{N}^0 = I_{\mathrm{I}_\mathrm{N}}$
2. **For** iter : 1 to $\mathrm{Iteration}_{\max}$
   (a) **For** k=1 to N
   - $\boldsymbol{U} = \boldsymbol{X} \times_1 W_1^{\mathrm{iter}-1} \ldots \times_{\mathrm{k}-1} W_{\mathrm{k}-1}^{\mathrm{iter}-1} \times_{\mathrm{k}+1} W_{\mathrm{k}+1}^{\mathrm{iter}-1} \ldots \times_\mathrm{N} W_\mathrm{N}^{\mathrm{iter}-1}$
   - $U^\mathrm{k} \Leftarrow_\mathrm{k} \boldsymbol{U}$
   - $\boldsymbol{Y} = \boldsymbol{Z} \times_1 W_1^{\mathrm{iter}-1} \ldots \times_{\mathrm{k}-1} W_{\mathrm{k}-1}^{\mathrm{iter}-1} \times_{\mathrm{k}+1} W_{\mathrm{k}+1}^{\mathrm{iter}-1} \ldots \times_\mathrm{N} W_\mathrm{N}^{\mathrm{iter}-1}$
   - $Y^\mathrm{k} \Leftarrow_\mathrm{k} \boldsymbol{Y}$
   - $\mathrm{n_I}\Sigma_I = \sum_{\mathrm{p}=1}^{\prod_{\mathrm{o}\neq\mathrm{k}} \mathrm{I_o}} \mathrm{n_I}\Sigma_I^\mathrm{p}, \mathrm{n_I}\Sigma_I^\mathrm{p} = \tilde{U}^{\mathrm{k,p}}(\tilde{U}^{\mathrm{k,p}})^T + \tilde{Y}^{\mathrm{k,p}}(\tilde{Y}^{\mathrm{k,p}})^T - S^{\mathrm{k,p}}(R^{\mathrm{k,p}})^T - R^{\mathrm{k,p}}(S^{\mathrm{k,p}})^T$
   - $\mathrm{n_E}\Sigma_E = \sum_{\mathrm{p}=1}^{\prod_{\mathrm{o}\neq\mathrm{k}} \mathrm{I_o}} \mathrm{n_E}\Sigma_E^\mathrm{p}, \mathrm{n_E}\Sigma_E^\mathrm{p} = \mathrm{m}U^{\mathrm{k,p}}(U^{\mathrm{k,p}})^T + \mathrm{n}Y^{\mathrm{k,p}}(Y^{\mathrm{k,p}})^T - \mathbf{s}^{\mathrm{k,p}}(\mathbf{r}^{\mathrm{k,p}})^T - \mathbf{r}^{\mathrm{k,p}}(\mathbf{s}^{\mathrm{k,p}})^T - \mathrm{n_I}\Sigma_I^\mathrm{p}$
   - Compute $\Sigma_E^\mathrm{k} W_\mathrm{k}^{\mathrm{iter}} = \Lambda_\mathrm{k} \Sigma_\mathrm{I}^\mathrm{k} W_\mathrm{k}^{\mathrm{iter}}$, obtain $W_\mathrm{k}^{\mathrm{iter}}$.

   (b) **If** iter > 2 and $\left\| W_\mathrm{k}^{\mathrm{iter}} - W_\mathrm{k}^{\mathrm{iter}-1} \right\| < \mathrm{I_k I_k} \epsilon, \mathrm{k} = 1, \cdots, \mathrm{N}$, break;
3. Sort the $\acute{\mathrm{I}}_\mathrm{N}$ eigenvectors $W_\mathrm{k}^{\mathrm{iter}} \in \Re^{\mathrm{I}_\mathrm{k} \times \acute{\mathrm{I}}_\mathrm{k}}$ according to Λ_k in decreasing order, $\mathrm{k} = 1, \cdots, \mathrm{N}$.

Journal Pre-proof

Journal Pre-proof

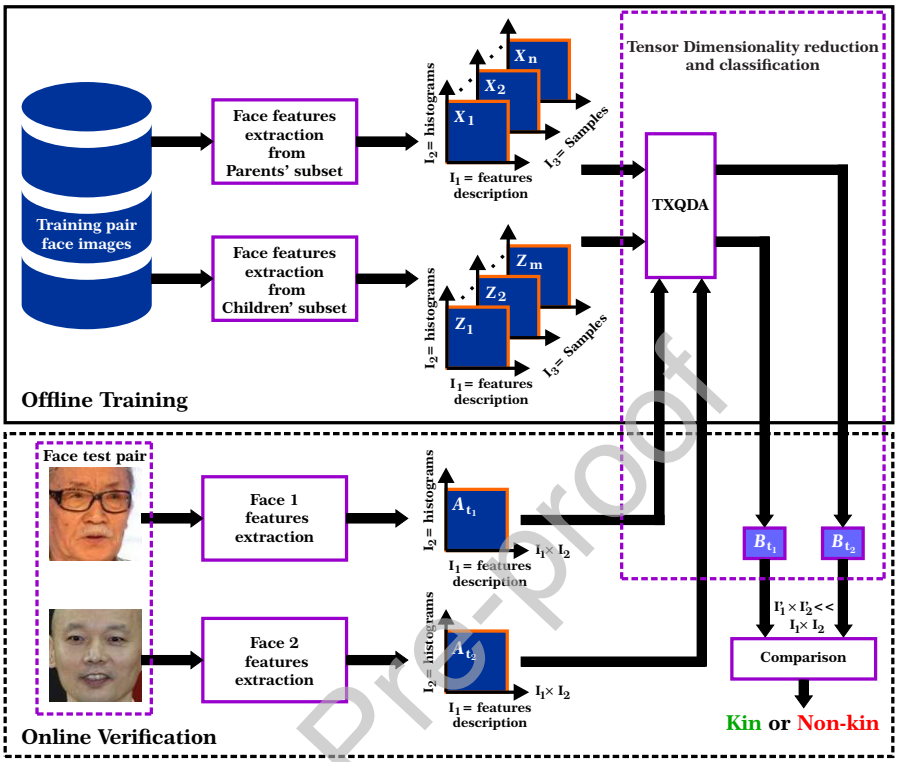


Figure 1: Block diagram of the proposed face pair matching system.

4.1. Feature extraction

To describe face images, we extract two popular local texture descriptors: the Binarized Statistical Image Feature (BSIF) [15] and the Local Phase Quantization (LPQ) [14]. To increase the verification rate, we extract the two descriptors at multiple scales by varying the values of their parameters, i.e., W the filter size of BSIF; M the window size of LPQ.

4.2. Tensor Design

In the offline (training) stage, the optimal multilinear projection matrices are estimated, and in online (test) stage, new samples are projected by these tensors and matched. The training 3^{rd} order tensors $\boldsymbol{X}, \boldsymbol{Z} \in \Re^{I_1 \times I_2 \times I_3}$ is constructed using the histograms of different local descriptors extracted from the training

Journal Pre-proof

face images. The three modes of the tensors $\boldsymbol{X}$ and $\boldsymbol{Z}$ are defined as follows: I_1 corresponds to the local descriptors extracted at different scales, I_2 represents the histograms, and I_3 face samples in the database .

The input tensors $\boldsymbol{X}$ and $\boldsymbol{Z}$ are reduced according to I_1 and I_2 modes and projected into another subspace based on the proposed TXQDA method. Then, we obtain a reduced tensor with $\acute{I_1} \times \acute{I_2} \ll I_1 \times I_2$.

The training data of TXQDA method includes the match pairs (positive pairs) only. The two tensors ($\boldsymbol{X}$ and $\boldsymbol{Z}$) are used to compute the covariance matrix of the intrapersonal variations Σ_I and the covariance matrix of the intrapersonal variations Σ_E of the TXQDA method.

In the test phase, each face pair passes the same steps of feature extraction as in the training phase, then projected in the tensor dimensionality reduction and classification (TXQDA). Finally, the cosine similarity is used to check whether the pair of reduced features matches (belonging to the same family) or not.

4.3. Matching

To compare between two faces pair, we used the reduced features projected through the TXQDA space which are concatenated to form one feature vector. Then, we applied cosine similarity [38] for each pair test of the two face images, so a match score is done. After discriminant analysis methods, the use of cosine similarity distance has an advantage which comes from its connection to the Bayes decision rule [39]. Cosine similarity between two vectors $\mathbf{b}_{t_1}$ and $\mathbf{b}_{t_2}$ is defined as the following:

$$\cos(\mathbf{b}_{t_1}, \mathbf{b}_{t_2}) = \frac{\mathbf{b}_{t_1}^{T}.\mathbf{b}_{t_2}}{\|\mathbf{b}_{t_1}\|.\|\mathbf{b}_{t_2}\|} \tag{9}$$

Where $\|.\|$ is the Euclidean norm. A high value of the produced score means a high probability that $\mathbf{b}_{t_1}$ and $\mathbf{b}_{t_2}$ belong to the same family.

5. Experiments

In this section, we perform a number of experiments to evaluate the proposed kinship verification system and investigate the strengths of the Tensor

Journal Pre-proof

representation. Firstly, we present the benchmark databases used in our experiments. Then, we discuss the parameter settings used for kinship verification. Finally, we provide and discuss the results and compare them with those of the state of the art.

5.1. Benchmark databases

To evaluate the performance of the proposed kinship verification approach, we considered five kinship databases: Cornell KinFace database, UB KinFace database, TSKinFace database, KinFaceW-II database and FIW database. These databases consist of four kinds of parent-child relationships. The face images are with various ages and ethnicities, and captured under uncontrolled environments and no restriction in terms of pose.

Cornell KinFace database [40] consists of 143 pairs of parents and children images gathered from the web. There are 286 cropped frontal face images of size 100×100 pixels. Most of the images were taken from Google Images. To ensure that the facial extracted characteristics are in high quality, only frontal face images with a neutral facial expression are chosen. We note that, 7 families are taken out of the original database which consists of 150 families for privacy issues.

UB KinFace database [41] includes 600 images of 400 people which are divided into 200 pairs of child-young parent (set 1) and 200 pairs of child-old parent (set 2). These two sets of pairs are used to enhance, test, and evaluate kinship verification algorithms. Most of images in the database are real-world combinations of public figures (celebrities and politicians) from Internet. It is the first database that comprises all children, young parents and old parents for the purpose of kinship verification.

TSKinFace database [1] consits of two types of tri-subject kinship relations which are: Father-Mother-Daughter (FM-D) and Father-Mother-Son (FM-S). The FM-D contains 502 relations and FM-S has 513 relations (4060 face images). These images are from public figures gathered from the Internet. The face images are cropped using the position of eyes into 64×64 pixels res-

Journal Pre-proof

olution. For fair comparison, we restructured the database by separating the group of Father-Mother-Daughter into two groups Father-Daughter and Mother-Daughter kinship relations, and the group of FatherMother-Son into two groups Father-Son and Mother-Son kinship relations.

Kinship Face in the Wild database (KinFaceW-II) [2] gathered through Internet research, including some public figures with their parents and/or children. In the KinFaceW-II database, each kinship relation type contains 250 pairs. In total 2000 face images for KinFaceW-II.

FIW database [42] we considered the largest FIW kinship database using: four relations, Grandfather-Granddaughter (GF-GD), Grandfather-Grandson (GF-GS), Grandmother-Granddaughter (GM-GD) and Grandmother-Grandson (GM-GS) face subsets. In GF-GD subset, there are 7,078 pairs of images for positive and negative relations. In GF-GS subset, there are 4,830 pairs of images for positive and negative relations. In GM-GD subset, there are 6,512 pairs of images for positive and negative relations. In GM-GS subset, there are 4,614 pairs of images for positive and negative relations.

5.2. Parameter Settings

Our approach's performance is evaluated based on the same experimental protocol found in the literature [2, 3, 43], in which five-fold cross-validation for kin verification is performed by keeping the same number of pairs images for each fold. This protocol is considered to make sure that our results are directly comparable to the state of the art. The negative pairs for kinship are generated randomly such that each image appears only once in the test set. The number of positive pairs and negative pairs is the same in the test stage.

To mitigate the effect of face normalization and to be consistent with several previous works, including [1, 2, 3, 4, 44, 5, 6, 7, 9, 10], all feature descriptions are extracted from face images that are aligned and cropped using the position of the eyes to 64×64 pixels.

Regarding feature extraction, we use eight filters with different sizes W = {3, 5, 7, 9, 11, 13, 15, 17} in the Multi-Scale Binarized Statistical Image Features

Journal Pre-proof

(MSBSIF). In the Multi-Scale Local Phase Quantization (MSLPQ), the window size is M = {3, 5, 7, 9, 11, 13, 15, 17} . Every face image is subdivided into 16 blocks, each of size 16×16 pixels. We use histograms of 256 bins to aggregate the local features extracted from each block.

5.3. Results and discussion

In this subsection, we introduce and discuss the results of the proposed approach based on third order tensor representation. Moreover, all the experiments were done for the original linear approach XQDA, which works as a baseline for evaluating the proposed TXQDA method. Furthermore, the performances of the local descriptors MSLPQ and MSBSIF are separately examined as well as their fusion. In the linear case, feature level fusion is made by concatenating vectors from different scales for each face descriptors. For the proposed multilinear case, fusion is performed based on tensor, where vectors of different scales of two descriptors, LPQ and BSIF, are stacked in the second mode of the tensor.

Tables 1, 2, 3, 4 and 5 show kinship verification accuracy from different descriptors and their fusion using the proposed TXQDA method compared with linear XQDA method on Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases, respectively. We remarked that the performance is improved with variation between 5% and 9%. Moreover, the proposed TXQDA method stacked the features in the second tensor mode to provide the maximum extraction of information. Consequently, many more feature dimensions are available in TXQDA than in XQDA. Furthermore, XQDA is theoretically limited by the number of classes in the data, whereas the TXQDA is not. It is also noticeable that the best results are obtained by ($\text{MSLPQ}_{(3+5+7+9+11)}$ + $\text{MSBSIF}_{(3+5+7+9+11)}$) description.

In Table 3, the tri-subject kinship verification is performed by Logistic Regression (LR) [45] score fusion method as used in [46]. The scores of Father-Son and Mother-Son are fused to generate FM-S scores for tri-subject matching. The scores of Father-Daughter and Mother-Daughter are fused to generate FM-D

Journal Pre-proof

scores for tri-subject matching.

As shown in Table 2 our TXQDA method processes the age difference factor and benefits from the child-young parent set (Set 1) in which this set is gathered to lighten the age difference shown in the child-old parent set (Set 2). XQDA method does not benefit from the child-young parent set and this demonstrates that the XQDA method neglected to process the age difference factor. Moreover, the multifactor structure (kinship, gender, age, expression, illumination and pose) was analyzed and separated from different tensor dimensions of the proposed TXQDA method.

Table 6 shows kinship verification accuracy of the proposed TXQDA method compared with linear XQDA method and three other methods (i.e. Neighborhood Repulsed Metric Learning [2] (NRML) method, Side-Information based Linear Discriminant analysis [47] (SILD) method and Multilinear Side-Information based Discriminant Analysis [48] (MSIDA) method) using (MSLPQ+MSBSIF) features description on Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases.

Moreover, the NRML and SILD methods, which are the most used methods for kinship verification, give lower performance than the XQDA and TXQDA methods in all cases. The performance becomes clear, significant and better by using the viewpoint changes methods compared with NRML and SILD methods.

Furthermore, Cornell KinFace and UB KinFace databases are gathered with mixture of four kin relations, which make hard learning of kin relations from different persons with different gender and with high significant age difference. Our proposed Tensor cross-view based method shows the superiority with a large margin in results compared with NRML and SILD methods on Cornell KinFace and UB KinFace. On Cornell KinFace database, TXQDA method performs with 9% of performance better than XQDA method, 9.42% of performance better than MSIDA, 18% of performance better than SILD and 17.5 % of performance better than NRML. On UB KinFace database, TXQDA method works with a performance of 9% better than XQDA method, a performance of 9.48% better than MSIDA, a performance of 24% better than SILD and a performance of 21

Journal Pre-proof

% better than NRML.

This demonstrates that the proposed Tensor XQDA works well on the difficult cases, where the viewpoint changes exist under the mixture of four kin relations (i.e. Father-Daughter and Mother-Son are a face images pairs that include two different person with different gender, and Father-Son and Mother-Daughter are a face images pairs that include two different person with same gender). Furthermore, TXQDA method works with a performance of 4.4% better than XQDA method, a performance of 4.83% better than MSIDA, a performance of 6.8% better than SILD and a performance of 7.3 % better than NRML on the TSKinFace database. Our TXQDA method works with a performance of 5.50% better than XQDA method, a performance of 5.00% better than MSIDA, a performance of 10.65% better than SILD and a performance of 14.20% better than NRML on the KinFaceW-II database. Moreover, our TXQDA method works with a performance of 7.93% better than XQDA method, a performance of 9.98% better than MSIDA, a performance of 11.05% better than SILD and a performance of 9.61% better than NRML on the FIW database.

Besides, our results show that the multifactor structure belonging to kinship, gender, age, expression, illumination and pose is taken into consideration in our TXQDA method to a large extent than the XQDA method. It is remarkable that SILD method used the positive and negative pairs in the training stage unlike the proposed TXQDA which obtained better performances than SILD by using only the positive pairs. Our TXQDA presented the face image as a matrix, where the face feature descriptions are stacked in a tensor mode providing the maximum extraction of information. Unlike in the case of XQDA method, the feature vectors are purely concatenated neglecting the natural structure of data. As shown in our results, the tensor is an elegant way of presenting and fusing data.

Journal Pre-proof

Journal Pre-proof

Table 1: The mean accuracy (%) of kinship verification for TXQDA and XQDA using different MSLPQ and MSBSIF scales and their fusion on the Cornell KinFace database.

Descriptor	XQDA	TXQDA
	Mean	Mean
$MSLPQ_{(3+5+7)}$	83.93	91.71
$MSLPQ_{(5+7+9)}$	81.07	92.71
$MSLPQ_{(9+11+13)}$	77.17	90.59
$MSLPQ_{(13+15+17)}$	76.51	89.89
$MSLPQ_{(3+5+7+9)}$	83.12	92.38
$MSLPQ_{(3+5+7+9+11)}$	83.83	92.74
$MSBSIF_{(3+5+7)}$	81.61	92.66
$MSBSIF_{(5+7+9)}$	79.60	92.32
$MSBSIF_{(9+11+13)}$	77.15	92.28
$MSBSIF_{(13+15+17)}$	79.65	89.85
$MSBSIF_{(3+5+7+9)}$	80.97	92.00
$MSBSIF_{(3+5+7+9+11)}$	80.98	92.33
$MSLPQ_{(3+5+7+9+11)}$ + $MSBSIF_{(3+5+7+9+11)}$	**84.10**	**93.04**

Journal Pre-proof

Table 2: The mean accuracy (%) of kinship verification for TXQDA and XQDA using different MSLPQ and MSBSIF scales and their fusion on the UB KinFace database.

Descriptor	XQDA			TXQDA		
	Set 1	Set 2	Mean	Set 1	Set 2	Mean
$MSLPQ_{(3+5+7)}$	76.74	80.21	78.48	91.28	90.53	90.91
$MSLPQ_{(5+7+9)}$	78.50	80.96	79.73	91.03	89.27	90.15
$MSLPQ_{(9+11+13)}$	74.71	76.19	75.45	90.25	88.29	89.27
$MSLPQ_{(13+15+17)}$	75.47	75.66	75.57	87.00	88.04	87.52
$MSLPQ_{(3+5+7+9)}$	79.22	80.19	79.71	91.28	90.02	90.65
$MSLPQ_{(3+5+7+9+11)}$	78.46	80.93	79.70	91.76	90.26	91.01
$MSBSIF_{(3+5+7)}$	77.00	76.73	76.87	91.02	90.77	90.90
$MSBSIF_{(5+7+9)}$	79.69	78.90	79.30	91.26	91.28	91.27
$MSBSIF_{(9+11+13)}$	79.23	80.47	79.85	91.28	90.76	91.02
$MSBSIF_{(13+15+17)}$	77.39	73.43	75.41	89.78	89.02	89.40
$MSBSIF_{(3+5+7+9)}$	82.25	79.70	80.98	91.25	91.03	91.14
$MSBSIF_{(3+5+7+9+11)}$	79.45	80.41	79.93	91.51	90.77	91.14
$MSLPQ_{(3+5+7+9+11)}$ + $MSBSIF_{(3+5+7+9+11)}$	82.24	82.92	**82.58**	92.03	91.02	**91.53**

Table 3: The mean accuracy (%) of kinship verification for TXQDA and XQDA using different MSLPQ and MSBSIF scales and their fusion on the TSKinFace database.

Descriptor	XQDA							TXQDA						
	F-S	F-D	M-S	M-D	Mean	FM-S	FM-D	F-S	F-D	M-S	M-D	Mean	FM-S	FM-D
$MSLPQ_{(3+5+7)}$	86.84	83.47	83.33	85.06	84.68	87.41	85.46	87.18	88.42	88.54	88.69	88.21	92.82	94.45
$MSLPQ_{(5+7+9)}$	83.63	81.78	83.33	84.66	83.35	86.64	85.36	84.27	85.05	85.73	86.11	85.29	91.26	91.77
$MSLPQ_{(9+11+13)}$	81.48	79.18	79.53	81.37	80.39	85.20	84.86	82.23	81.78	84.08	83.13	82.81	90.58	90.18
$MSLPQ_{(13+15+17)}$	80.90	78.69	78.95	78.28	79.21	85.38	83.67	81.55	80.89	83.69	82.63	82.19	90.19	87.89
$MSLPQ_{(3+5+7+9)}$	85.77	84.66	84.99	86.35	85.44	87.31	86.15	87.09	87.72	88.16	89.09	88.02	92.91	94.45
$MSLPQ_{(3+5+7+9+11)}$	86.84	84.06	83.72	84.76	84.85	87.22	85.95	88.35	88.61	88.25	89.28	88.62	93.01	94.44
$MSBSIF_{(3+5+7)}$	83.14	82.76	81.77	85.15	83.21	85.78	85.25	88.06	88.32	88.54	89.68	88.65	93.98	94.34
$MSBSIF_{(5+7+9)}$	84.32	82.57	81.97	84.56	83.36	86.74	86.85	86.41	88.12	88.54	88.10	87.79	93.69	94.74
$MSBSIF_{(9+11+13)}$	82.95	79.99	80.51	82.67	81.53	84.79	84.36	85.15	85.45	86.80	86.01	85.85	92.52	91.96
$MSBSIF_{(13+15+17)}$	82.27	75.90	79.14	79.29	79.15	84.99	85.45	84.17	83.37	86.70	85.11	84.84	91.75	91.46
$MSBSIF_{(3+5+7+9)}$	84.90	84.26	82.26	86.15	84.39	87.52	87.35	89.03	88.32	89.13	89.48	88.99	94.08	94.44
$MSBSIF_{(3+5+7+9+11)}$	85.29	83.06	84.02	86.35	84.68	87.32	87.35	89.03	88.12	89.13	89.38	88.92	94.08	94.34
$MSLPQ_{(3+5+7+9+11)}$ + $MSBSIF_{(3+5+7+9+11)}$	87.04	85.26	84.40	86.85	**85.89**	**88.02**	**88.35**	89.32	90.69	90.29	90.97	**90.32**	**94.85**	**95.63**

Journal Pre-proof

Journal Pre-proof

Table 4: The mean accuracy (%) of kinship verification for TXQDA and XQDA using different MSLPQ and MSBSIF scales and their fusion on the KinFaceW-II database.

Descriptor	XQDA					TXQDA				
	F-S	F-D	M-S	M-D	Mean	F-S	F-D	M-S	M-D	Mean
$MSLPQ_{(3+5+7)}$	83.80	77.40	79.00	80.20	80.10	88.00	83.20	83.60	83.60	84.60
$MSLPQ_{(5+7+9)}$	83.80	78.20	79.60	77.80	79.85	89.20	83.40	84.00	84.20	85.20
$MSLPQ_{(9+11+13)}$	81.80	79.00	77.80	78.40	79.25	87.00	81.40	83.80	85.20	84.35
$MSLPQ_{(13+15+17)}$	80.80	77.60	75.40	78.20	78.00	89.00	80.80	81.80	83.60	83.80
$MSLPQ_{(3+5+7+9)}$	84.20	77.60	80.20	80.00	80.50	89.20	83.80	83.00	84.80	85.20
$MSLPQ_{(3+5+7+9+11)}$	84.00	78.80	80.40	80.00	80.80	89.40	83.40	83.60	85.00	85.35
$MSBSIF_{(3+5+7)}$	83.60	78.60	79.20	79.60	80.25	86.60	85.20	82.20	83.20	84.30
$MSBSIF_{(5+7+9)}$	84.40	79.40	78.80	77.80	80.10	88.00	84.20	83.40	82.00	84.10
$MSBSIF_{(9+11+13)}$	83.40	78.60	77.40	77.20	79.15	86.60	82.20	81.20	82.00	83.00
$MSBSIF_{(13+15+17)}$	82.00	76.00	77.20	76.60	77.95	86.40	81.20	82.80	81.60	83.00
$MSBSIF_{(3+5+7+9)}$	84.40	80.20	79.20	79.60	80.85	87.60	85.20	81.60	83.80	84.55
$MSBSIF_{(3+5+7+9+11)}$	84.60	79.20	78.80	79.60	80.55	88.00	85.80	82.20	85.00	85.25
$MSLPQ_{(3+5+7+9+11)}$ + $MSBSIF_{(3+5+7+9+11)}$	85.00	80.60	80.60	80.40	**81.65**	90.20	86.40	85.60	86.40	**87.15**

Table 5: The mean accuracy (%) of kinship verification for TXQDA and XQDA using different MSLPQ and MSBSIF scales and their fusion on the four grandparent-grandchild subsets of FIW database.

Descriptor	XQDA					TXQDA				
	GF-GD	GF-GS	GM-GD	GM-GS	Mean	GF-GD	GF-GS	GM-GD	GM-GS	Mean
$MSLPQ_{(3+5+7)}$	55.91	57.92	58.08	57.59	57.38	65.57	63.92	63.85	65.08	64.61
$MSLPQ_{(5+7+9)}$	55.73	58.07	57.78	57.87	57.36	65.54	64.81	63.96	64.62	64.73
$MSLPQ_{(9+11+13)}$	55.98	58.30	57.85	57.16	57.32	66.11	64.23	63.00	64.54	64.47
$MSLPQ_{(13+15+17)}$	56.07	58.06	57.22	56.70	57.01	65.49	65.18	62.65	64.14	64.37
$MSLPQ_{(3+5+7+9)}$	56.15	58.29	57.98	58.54	57.74	65.89	64.43	63.82	64.71	64.71
$MSLPQ_{(3+5+7+9+11)}$	56.36	58.90	58.27	58.15	57.92	65.32	63.79	63.49	64.15	64.19
$MSBSIF_{(3+5+7)}$	55.85	59.06	58.46	57.26	57.66	65.35	64.50	64.32	64.61	64.70
$MSBSIF_{(5+7+9)}$	55.58	58.03	59.27	57.06	57.49	65.25	64.21	63.52	64.45	64.36
$MSBSIF_{(9+11+13)}$	56.30	56.80	58.44	56.73	57.07	66.00	65.04	64.74	64.95	65.18
$MSBSIF_{(13+15+17)}$	56.32	58.01	57.81	56.53	57.17	65.46	64.18	63.24	64.75	64.41
$MSBSIF_{(3+5+7+9)}$	56.04	58.61	59.00	57.60	57.81	65.20	64.59	63.15	64.43	64.34
$MSBSIF_{(3+5+7+9+11)}$	55.93	58.43	59.35	57.70	57.85	65.26	64.22	63.78	65.33	64.65
$MSLPQ_{(3+5+7+9+11)}$ + $MSBSIF_{(3+5+7+9+11)}$	56.04	59.23	59.00	58.12	**58.10**	66.43	66.79	65.24	65.67	**66.03**

To analyze the performance of different kinship relations, we plot in Figures 2, 3, 4 and 5 the ROC curves of different methods (NRML, SILD, XQDA,

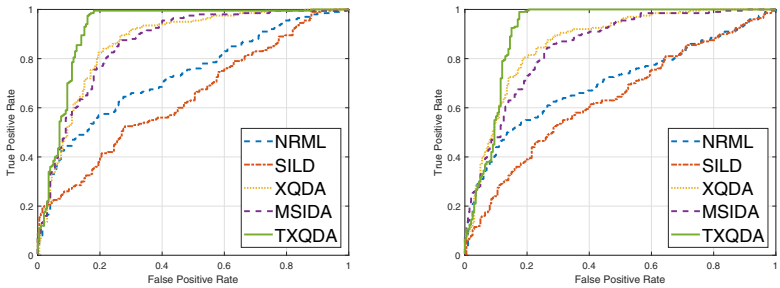
Journal Pre-proof

Table 6: Comparison verification accuracy (%) of the proposed TXQDA with XQDA, NRML, SILD and MSIDA methods using MSLPQ+MSBSIF features description on the Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases.

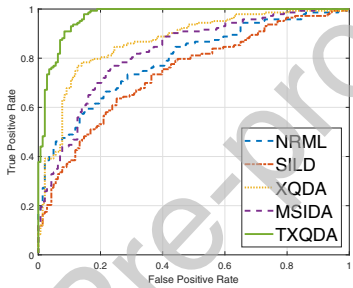
Method	Cornell KinFace	UB KinFace	TSKinFace			KinFaceW-II	FIW
	Mean	Mean	Mean	FM-S	FM-D	Mean	Mean
NRML	75.52	70.55	80.83	84.26	85.53	72.95	56.42
SILD	71.38	67.36	83.46	86.44	87.82	76.50	54.98
XQDA	84.10	82.58	85.89	88.02	88.35	81.65	58.10
MSIDA	83.62	82.05	85.49	91.26	91.27	82.15	56.05
TXQDA	**93.04**	**91.53**	**90.32**	**94.85**	**95.63**	**87.15**	**66.03**

MSIDA and TXQDA) using the best performing features (MSLPQ+MSBSIF) on UB KinFace set 1, UB KinFace set 2, Cornell KinFace, TSKinFace, KinFaceW-II and FIW databases, respectively.

Journal Pre-proof



(a) (b)



(c)

Figure 2: ROC curves of different methods (NRML, SILD, XQDA, MSIDA and TXQDA) using the best performing features (MSLPQ+MSBSIF) obtained on (a) UB KinFace set 1, (b) UB KinFace set 2, (c) Cornell KinFace databases, respectively.

Journal Pre-proof

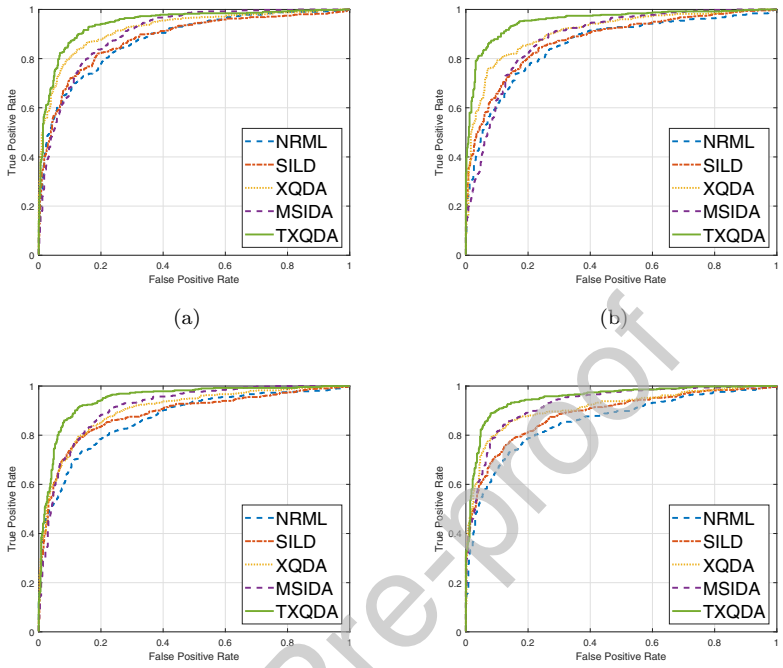


Figure 3: ROC curves of different methods (NRML, SILD, XQDA, MSIDA and TXQDA) using the best performing features (MSLPQ+MSBSIF) on TSKinFace database obtained on (a) F-S set, (b) F-D set, (c) M-S set and (d) M-D set, respectively.

Journal Pre-proof

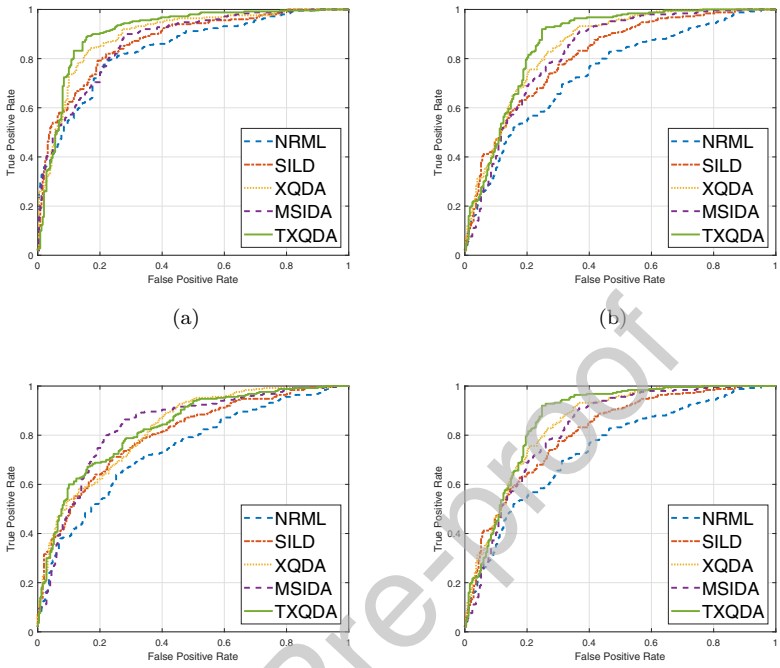


Figure 4: ROC curves of different methods (NRML, SILD, XQDA, MSIDA and TXQDA) using the best performing features (MSLPQ+MSBSIF) on KinFaceW-II database obtained on (a) F-S set, (b) F-D set, (c) M-S set and (d) M-D set, respectively.

Journal Pre-proof

Journal Pre-proof

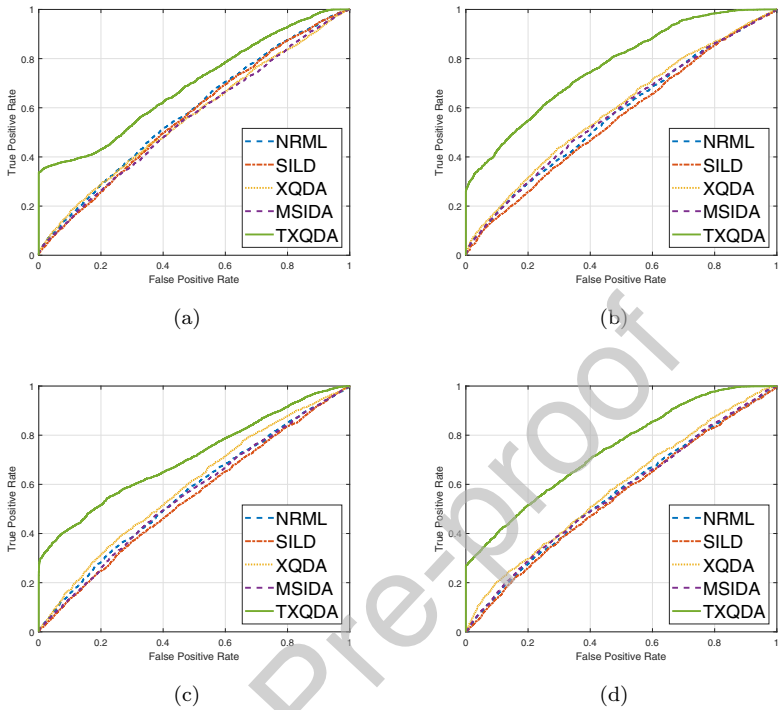


Figure 5: ROC curves of different methods (NRML, SILD, XQDA, MSIDA and TXQDA) using the best performing features (MSLPQ+MSBSIF) on FIW database obtained on (a) GF-GD set, (b) GF-GS set, (c) GM-GD set and (d) GM-GS set, respectively.

5.4. The robustness's evaluation of the proposed TXQDA method

In this subsection, we tested the robustness of the proposed tensor method on TSKinFace database through additive noise and degradation of test set. For the clarification, we express the interference of face recognition $\eta = \eta_f + \eta_q$ [49], where η_f indicates facial variations such as kinship, expression, illumination, misalignment and age, and η_q indicates the image variation due to sensor or coding-related issues, such as Gaussian noise, blur, compression, and low resolution. Most of the studies on the TSKinFace database focused only on the effect of η_f, whereas our extended experiments study both the pure effect of η_q and the superposed interference of $\eta_f + \eta_q$. For an inclusive study, we generate four types of noise or degradations that are most common in real-world systems

Journal Pre-proof

but that have not appeared in the standard database. Specifically, we generate the following versions of test sets: 1) three levels of Gaussian noise. The images are normalized in the range of (0; 1), and then we apply additive Gaussian noise with zero mean and standard derivations of $\sigma = 0.01; 0.02; 0.03$; 2) three different Gaussian blur test sets using a Gaussian kernel of size 10×10 with $\sigma = \{1; 3/2; 2\}$; 3) three different compressed images using MATLABs JPEG codec of quality 60, 45 and 30; and 4) three different low-resolution sets of test images by first downsampling the images by ratios of 2, 3, and 4 and then interpolating them to the original resolution by the nearest method in MATLAB. Example test images are shown in Fig. 6, and as shown in this figure, these degraded faces are recognizable by humans and are very common in real-world surveillance scenarios. Therefore, it is important to study how the accuracy of the metric learning methods change under these degradations. Table 7 shows that TXQDA and XQDA show much better robustness than the other methods under image blur, noise, compression, and reduced resolution. However, our Tensor XQDA preserves the data structure and extracts more discriminative information from degraded face images compared with XQDA method.

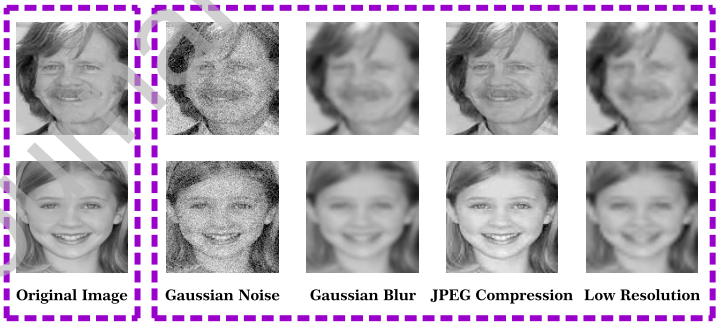


Figure 6: Examples of original and degraded images used in our extended TSKinFace evaluation. The last four columns correspond to the most severe degrees of Gaussian noise, Gaussian blur, JPEG compression, and reduced resolution applied on the test images.

Journal Pre-proof

Table 7: Comparative verification rates (%) of extended TSKinFace evaluation on the robustness to the four types of common degradations. Accuracy loss of each degradation degree on each test set is reported in detail.

Method	Relation Type	Basic Accuracy [1]	Gaussian Blur			Gaussian Noise			JPEG Compression			Reduced Resolution			Summarized Accuracy [2]
			1	3/2	2	0.01	0.02	0.03	60	45	30	1/2	1/3	1/4	
NRML	FS	81.00	-7.02	-7.81	-12.68	-6.93	-8.49	-14.33	-5.95	-6.63	-7.22	-6.73	-7.12	-12.97	72.34 (-8.66)
	FD	79.88	-6.27	-8.47	-13.65	-7.67	-9.66	-13.45	-7.07	-9.16	-12.85	-6.57	-8.67	-13.94	70.09 (-9.79)
	MS	81.48	-8.67	-10.92	-14.13	-9.25	-11.11	-13.15	-9.06	-10.62	-13.64	-8.48	-10.43	-13.94	70.36 (-11.12)
	MD	80.97	-5.38	-9.37	-13.85	-5.38	-8.47	-13.35	-5.38	-8.67	-13.15	-5.38	-8.67	-13.95	71.72 (-9.25)
	Mean	80.83	-6.83	-9.14	-13.57	-7.3	-9.43	-13.57	-6.86	-8.77	-11.71	-6.79	-8.72	-13.7	71.13 (-9.7)
SILD	FS	84.60	-9.25	-10.81	-12.76	-9.35	-10.23	-13.35	-6.43	-8.57	-9.06	-9.06	-9.45	-11.20	74.64 (-9.96)
	FD	81.76	-9.05	-11.64	-12.94	-11.64	-11.84	-12.44	-8.65	-8.85	-9.25	-8.65	-9.35	-10.94	71.32 (-10.44)
	MS	83.72	-10.52	-12.86	-16.46	-10.71	-12.57	-15.68	-10.32	-11.01	-11.20	-11.30	-11.98	-13.44	71.38 (-12.34)
	MD	83.76	-9.17	-11.16	-15.04	-8.88	-10.66	-14.55	-9.97	-11.06	-14.15	-10.07	-10.87	-11.56	72.28 (-11.48)
	Mean	83.46	-9.50	-11.62	-14.30	-10.15	-11.32	-14.01	-8.84	-9.87	-10.91	-9.92	-10.41	-11.79	72.41 (-11.05)
XQDA	FS	87.04	-6.73	-8.87	-10.72	-5.85	-7.12	-8.87	-3.90	-7.12	-7.99	-6.14	-8.87	-10.73	79.30 (-7.74)
	FD	85.26	-5.08	-7.77	-10.25	-4.08	-4.98	-7.07	-3.59	-4.78	-4.98	-3.98	-6.68	-10.16	79.14 (-6.12)
	MS	84.40	-4.09	-6.05	-8.68	-3.61	-4.48	-5.75	-3.12	-4.87	-5.27	-3.61	-5.65	-8.87	79.06 (-5.34)
	MD	86.85	-5.88	-7.27	-8.87	-5.68	-7.87	-7.97	-5.68	-7.27	-7.27	-6.28	-6.87	-9.16	79.68 (-7.17)
	Mean	85.89	-5.45	-7.49	-9.63	-4.81	-6.11	-7.42	-4.08	-6.01	-6.38	-5.00	-7.02	-9.73	79.30 (-6.59)
TXQDA	FS	89.32	-1.46	-4.56	-7.18	-0.87	-4.37	-6.80	-2.14	-3.01	-3.59	-2.62	-4.37	-7.09	85.32 (-4.00)
	FD	90.69	-2.27	-5.15	-9.70	-1.38	-4.95	-8.41	-3.07	-4.16	-4.55	-3.26	-5.74	-10.49	85.43 (-5.26)
	MS	90.29	-1.94	-5.63	-8.64	-0.78	-4.46	-7.57	-1.26	-3.69	-4.17	-2.62	-4.66	-7.86	85.85 (-4.44)
	MD	90.97	-1.68	-5.46	-8.53	-1.29	-5.65	-6.95	-1.59	-4.17	-5.46	-2.48	-4.96	-8.13	86.27 (-4.70)
	Mean	90.32	-1.84	-5.20	-8.51	-1.08	-4.86	-7.43	-2.02	-3.76	-4.62	-2.75	-4.93	-8.39	**85.72 (-4.60)**

To provide a comprehensive result, the verification accuracy across the three types of probe sets is reported.

[1] The verification accuracy on the original TSKinFace database.

[2] The verification accuracy across all types and all degrees of the tested noise and degradations.

5.5. Computational cost

We computed the computational time needed for the kinship verification of one pair face samples (Parent-Child) using different methods. The experiments were implemented using MATLAB 2014a on a PC with an Intel Core i7 2.00 GHz CPU and 8 GB of RAM. The feature extraction for the case of MSLPQ+MSBSIF description takes 0.104 s. In training stage (offline), the estimation of the projection matrices is performed only once. In the online phase, we evaluate the time cost needed by each method to project and match the test pair which is provided in Table 8 in ms. This table shows that the best performing variant, TXQDA, runs faster compared with all the other methods. Furthermore, we see that the time cost of projection and matching is negligible compared to the feature extraction time. The total time cost for our framework using MSLPQ+MSBSIF features and TXQDA method is about 0.109 s for both Cornell KinFace and UB KinFace databases, and 0.110 s from TSKinFace and

Journal Pre-proof

KinFaceW-II databases, and 0.115 s for FIW database.

Table 8: Time Cost (TC), in ms, taken by different methods for the projection of one pair of face images.

Database	Feature extraction	Projection and matching					All steps				
		NRML	SILD	XQDA	MSIDA	TXQDA	NRML	SILD	XQDA	MSIDA	TXQDA
Cornell KinFace	104.92	14.22	22.93	10.61	5.22	**4.54**	119.14	127.85	115.53	110.14	**109.46**
UB KinFace		15.17	11.91	7.06	5.73	**4.94**	120.09	116.83	111.98	110.65	**109.86**
TSKinFace		43.04	12.15	10.92	5.90	**5.88**	147.96	117.07	115.84	110.82	**110.80**
KinFaceW-II		21.17	11.99	11.03	6.32	**6.05**	126.09	116.91	115.95	111.24	**110.97**
FIW		97.27	90.11	78.96	21.16	**10.41**	102.19	195.03	183.88	126.08	**115.33**

5.6. Comparison with the results of the state of the art

The best kinship verification performances of our approach are achieved using two descriptors ($\text{MSLPQ}_{(3+5+7+9+11)}$ + $\text{MSBSIF}_{(3+5+7+9+11)}$) on Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases. For the XQDA (linear), verification rates of 84.10%, 82.58%, 85.89%, 81.65% and 58.10% are reported on Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases, respectively. For TXQDA (multilinear), verification rates of 93.04%, 91.53%, 90.32%, 87.15% and 66.03% are reported on Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases, respectively. These results are compared with the state of the art in Tables 9, 10, 11, 12 and 13. The comparison reveals that the proposed Tensor cross-view analysis based method outperforms the recent state of the art on the five databases, Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW. This demonstrates the effectiveness of using the cross-view methods in kinship verification topic.

TXQDA vs. KVRL-fcDBN [50]: In this comparison, we focus on two databases Cornell KinFace and UB KinFace. The work of Kohli *et al.* [50] was based on a deep learning approach (KVRL-fcDBN), where fcDBN algorithm was used to learn more than 600,000 outside face images and obtained 89.50 % on Cornell KinFace and 91.80 % on UB KinFace databases. However, in our work, we used Tensor cross-view metric learning method (i.e. TXQDA) to learn the provided data only (i.e. no outside data was used) and we achieved good performances, 93.04 % and 91.53 % on Cornell KinFace and UB KinFace

Journal Pre-proof

databases, respectively. Our TXQDA outperforms the KVRL-fcDBN approach on Cornell KinFace database and got a competitive performance on UB KinFace database.

TXQDA vs. DDML [7], DDMML [7] and MvDML [51]: In this comparison, we focus on TSKinFace and KinFaceW-II databases for monocular and multiple feature description. From mono-view feature description, the Discriminative Deep Metric Learning (DDML) [7] method used LPQ descriptor with face images with input size of 64×64. First they divide each face image into 4×4 non-overlapping blocks, where the size of each block is 16×16. Then, they extract a 256-bin LPQ histogram with window size of 3, 5 and 7 for each block respectively, and finally concatenate them to form a 12,288-D feature vector, which is the same used face features description in our work (i.e. $\text{MSLPQ}_{(3+5+7)}$). For bi-subject kinship verification, from the $\text{MSLPQ}_{(3+5+7)}$ features description, our TXQDA method outperforms the DDML method with about 8.28 % and 2.40 % on TSKinFace and KinFaceW-II databases, respectively. For tri-subject kinship verification, from the $\text{MSLPQ}_{(3+5+7)}$ features description, our TXQDA method outperforms the DDML method with about 9.12 % and 11.25 % on TSKinFace database for FM-S and FM-D relations, respectively. From multi-view feature description, we can see that our approach which using the MSLPQ+MSBSIF feature description, improves the performance with about 2.85 % and 6.95 % compared with DDMML and MvDML on KinFaceW-II database, respectively. Furthermore, our TXQDA method improves the performances with about 6 % compared with DDMML for bi-subject kinship verification on TSKinFace database and improves the performances with about 6.35 % and 8.53 % compared with DDMML for FM-S and FM-D tri-subject relations on TSKinFace database.

TXQDA vs. MSIDA [48]: From the five kinship databases, our TXQDA outperforms the MSIDA [48] method with about 9.42%, 9.48%, 4.83%, 5.00% and 9.98% on Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases, respectively. Furthermore, our TXQDA used only the positive pairs in the training step, unlike MSIDA do. The framework of Bessaoudi et al. [48]

Journal Pre-proof

needs the MPCA method step for the features dimension reduction before using the MSIDA method. Our TXQDA method deals with the face images directly, without the need of using the features dimension reduction step. Furthermore, our TXQDA method work well compared with MSIDA method when the data classes contain face images from different persons with very large age difference (i.e. the four grandparent-grandchild relations of FIW database).

Table 9: Performance comparisons (%) with state-of-the-art methods on Cornell KinFace database.

Method	Mean Accuracy (%)
Pictorial structure model [40]	70.67
Neighborhood repulsed metric learning [2]	69.50
Multiview neighborhood repulsed metric learning [2]	71.60
Discriminative multimetric learning [3]	73.50
Prototype discriminative feature learning [4]	71.90
MHDL3 - {HOG + Color + LPQ} [9]	76.60
Multiple kernel similarity metric [10]	81.70
Heterogeneous similarity learning [52]	68.40
Kinship metric learning [53]	81.40
Multilinear side-information based discriminant analysis [48]	86.59
Neighborhood repulsed metric learning [2] (Our)	75.52
Side-information based linear discriminant analysis [47] (Our)	71.38
Cross-view quadratic discriminant analysis [13] (Our)	84.10
Multilinear side-information based discriminant analysis [48] (Our)	83.62
TXQDA (Our)	**93.04**

Journal Pre-proof

Journal Pre-proof

Table 10: Table 8: Performance comparisons (%) with state-of-the-art methods on UB KinFace database.

Method	Mean Accuracy (%)
Transfer subspace learning [41]	68.50
Neighborhood repulsed metric learning [2]	65.60
Multiview neighborhood repulsed metric learning [2]	67.05
Discriminative multimetric learning [3]	72.25
Prototype discriminative feature learning [4]	67.30
Heterogeneous similarity learning [52]	56.20
PML-COV-S [54]	84.50
Kinship metric learning [53]	75.50
Multilinear side-information based discriminant analysis [48]	83.34
Neighborhood repulsed metric learning [2] (Our)	70.55
Side-information based linear discriminant analysis [47] (Our)	67.36
Cross-view quadratic discriminant analysis [13] (Our)	82.58
Multilinear side-information based discriminant analysis [48] (Our)	82.05
TXQDA (Our)	**91.53**

Table 11: Performance comparisons (%) with state-of-the-art methods on TSKinFace database.

Method	Mean Accuracy (%)	FM-S (%)	FM-D (%)
Relative symmetric bilinear model [1]	81.85	86.40	84.40
BSIF-IISV [5]	81.19	/	/
Discriminative deep multi-metric learning [7]	84.15	88.50	87.10
Multiple kernel similarity metric [10]	84.52	/	/
Multilinear side-information based discriminant analysis [48]	85.18	/	/
SILD+WCCN/LR [46]	88.59	90.94	91.23
Neighborhood repulsed metric learning [2] (Our)	80.83	84.26	85.53
Side-information based linear discriminant analysis [47] (Our)	83.46	86.44	87.82
Cross-view quadratic discriminant analysis [13] (Our)	85.89	88.02	88.35
Multilinear side-information based discriminant analysis [48] (Our)	85.49	91.26	91.27
TXQDA (Our)	**90.32**	**94.85**	**95.63**

Journal Pre-proof

Journal Pre-proof

Table 12: Performance comparisons (%) with state-of-the-art methods on KinFaceW-II database.

Method	Mean Accuracy (%)
Multi-view multi-task learning [8]	77.20
Discriminative deep multi-metric learning [7]	84.30
Multi-view deep metric learning [51]	80.20
Heterogeneous similarity learning [52]	70.40
Multiple kernel similarity metric [10]	84.30
Large-margin multi-metric learning [55]	80.00
Kinship metric learning [53]	85.70
SILD+WCCN/LR [46]	86.20
Neighborhood repulsed metric learning [2] (Our)	72.95
Side-information based linear discriminant analysis [47] (Our)	76.50
Cross-view quadratic discriminant analysis [13] (Our)	81.65
Multilinear side-information based discriminant analysis [48] (Our)	82.15
TXQDA (Our)	**87.15**

Table 13: Performance comparisons (%) with state-of-the-art methods on the four grandparent-grandchild relations from FIW database.

Method	Mean Accuracy (%)
ResNet+CF [42]	65.51
SphereFace [42]	65.60
ResNet+SDMLoss [56]	65.58
Neighborhood repulsed metric learning [2] (Our)	56.42
Side-information based linear discriminant analysis [47] (Our)	54.98
Cross-view quadratic discriminant analysis [13] (Our)	58.10
Multilinear side-information based discriminant analysis [48] (Our)	56.05
TXQDA (Our)	**66.03**

6. Conclusion

In this paper, we presented an effective approach based on tensor cross-view method to the problem of kinship verification. To achieve a low dimensional and discriminative tensor subspace, we extended XQDA to TXQDA, which operate on multilinear data. TXQDA finds multilinear projections of the tensor, where

Journal Pre-proof

the separation between data classes is enhanced. Furthermore, TXQDA was proposed to separate the multifactor structure of face images related to kinship, age, gender, expression, illumination and pose from different dimensions of the tensor. Therefore, TXQDA has many advantages as it, i) preserves data structure, ii) enlarges the margin between samples, iii) helps lightening the small sample size problem, and iv) reduces the computational cost. The experimental evaluation showed the superiority of our method. The best results of our approach are obtained by fusing histograms of two multiple scale local texture descriptors (MSLPQ+MSBSIF) projected with the proposed TXQDA method. These results outperform the state of the art on Cornell KinFace, UB KinFace, TSKinFace, KinFaceW-II and FIW databases. Furthermore, these results point out to the need of using cross-view methods for kinship verification. As future work, we plan to investigate higher tensor orders (> 3) for face representation with the proposed multilinear dimensionality reduction method.

References

[1] X. Qin, X. Tan, S. Chen, Tri-subject kinship verification: Understanding the core of a family, IEEE Transactions on Multimedia 17 (10) (2015) 1855–1867. doi:10.1109/TMM.2015.2461462.

[2] 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 (2) (2014) 331–345. doi:10.1109/TPAMI.2013.134.
URL http://dx.doi.org/10.1109/TPAMI.2013.134

[3] H. Yan, J. Lu, W. Deng, X. Zhou, Discriminative multimetric learning for kinship verification, IEEE Transactions on Information Forensics and Security 9 (7) (2014) 1169–1178. doi:10.1109/TIFS.2014.2327757.

[4] H. Yan, J. Lu, X. Zhou, Prototype-based discriminative feature learning for kinship verification, IEEE Transactions on Cybernetics 45 (11) (2015) 2535–2545. doi:10.1109/TCYB.2014.2376934.

Journal Pre-proof

[5] X. Wu, E. Boutellaa, M. B. Lpez, X. Feng, A. Hadid, On the usefulness of color for kinship verification from face images, in: 2016 IEEE International Workshop on Information Forensics and Security (WIFS), 2016, pp. 1–6. doi:10.1109/WIFS.2016.7823901.

[6] Z. Zhang, Y. Chen, V. Saligrama, Group membership prediction, in: 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 3916–3924. doi:10.1109/ICCV.2015.446.

[7] J. Lu, J. Hu, Y. P. Tan, Discriminative deep metric learning for face and kinship verification, IEEE Transactions on Image Processing 26 (9) (2017) 4269–4282. doi:10.1109/TIP.2017.2717505.

[8] X. Qin, X. Tan, S. Chen, Mixed bi-subject kinship verification via multi-view multi-task learning, Neurocomputing 214 (2016) 350 – 357. doi:https://doi.org/10.1016/j.neucom.2016.06.027.
URL http://www.sciencedirect.com/science/article/pii/S0925231216306658

[9] S. Mahpod, Y. Keller, Kinship verification using multiview hybrid distance learning, Computer Vision and Image Understanding 167 (2018) 28 – 36. doi:https://doi.org/10.1016/j.cviu.2017.12.003.
URL http://www.sciencedirect.com/science/article/pii/S107731421730228X

[10] Y.-G. Zhao, Z. Song, F. Zheng, L. Shao, Learning a multiple kernel similarity metric for kinship verification, Information Sciences 430-431 (2018) 247 – 260. doi:https://doi.org/10.1016/j.ins.2017.11.048.
URL http://www.sciencedirect.com/science/article/pii/S0020025516310416

[11] S. Liao, D. Yi, Z. Lei, R. Qin, S. Z. Li, Heterogeneous face recognition from local structures of normalized appearance, in: M. Tistarelli, M. S. Nixon (Eds.), Advances in Biometrics, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 209–218.

Journal Pre-proof

[12] D. Gray, H. Tao, Viewpoint invariant pedestrian recognition with an ensemble of localized features, in: D. Forsyth, P. Torr, A. Zisserman (Eds.), Computer Vision – ECCV 2008, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 262–275.

[13] S. Liao, Y. Hu, X. Zhu, S. Z. Li, Person Re-identification by Local Maximal Occurrence Representation and Metric Learning, ArXiv e-printsarXiv:1406.4216.

[14] V. Ojansivu, J. Heikkilä, Blur Insensitive Texture Classification Using Local Phase Quantization, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 236–243. doi:10.1007/978-3-540-69905-7_27.
URL http://dx.doi.org/10.1007/978-3-540-69905-7_27

[15] J. Kannala, E. Rahtu, Bsif: Binarized statistical image features, in: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), 2012, pp. 1363–1366.

[16] S. J. Wang, S. Yan, J. Yang, C. G. Zhou, X. Fu, A general exponential framework for dimensionality reduction, IEEE Transactions on Image Processing 23 (2) (2014) 920–930. doi:10.1109/TIP.2013.2297020.

[17] T. Zhang, B. Fang, Y. Y. Tang, Z. Shang, B. Xu, Generalized discriminant analysis: A matrix exponential approach, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 40 (1) (2010) 186–197. doi:10.1109/TSMCB.2009.2024759.

[18] S. Liu, Q. Ruan, C. Wang, G. An, Tensor rank one differential graph preserving analysis for facial expression recognition, Image Vision Comput. 30 (8) (2012) 535–545. doi:10.1016/j.imavis.2012.05.004.
URL http://dx.doi.org/10.1016/j.imavis.2012.05.004

[19] Y. Pang, S. Wang, Y. Yuan, Learning regularized lda by clustering, IEEE Transactions on Neural Networks and Learning Systems 25 (12) (2014) 2191–2201. doi:10.1109/TNNLS.2014.2306844.

Journal Pre-proof

[20] Z. Lai, Y. Xu, J. Yang, J. Tang, D. Zhang, Sparse tensor discriminant analysis, IEEE Transactions on Image Processing 22 (10) (2013) 3904–3915. doi:10.1109/TIP.2013.2264678.

[21] H. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, Mpca: Multilinear principal component analysis of tensor objects, IEEE Transactions on Neural Networks 19 (1) (2008) 18–39. doi:10.1109/TNN.2007.901277.

[22] H. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, Uncorrelated multilinear discriminant analysis with regularization and aggregation for tensor object recognition, IEEE Transactions on Neural Networks 20 (1) (2009) 103–123. doi:10.1109/TNN.2008.2004625.

[23] S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, H. J. Zhang, Multilinear discriminant analysis for face recognition, IEEE Transactions on Image Processing 16 (1) (2007) 212–220. doi:10.1109/TIP.2006.884929.

[24] H. Zhao, S. Sun, Sparse tensor embedding based multispectral face recognition, Neurocomputing 133 (2014) 427 – 436. doi:https://doi.org/10.1016/j.neucom.2013.12.019.
URL http://www.sciencedirect.com/science/article/pii/S0925231214000320

[25] K. Lee, H. Park, Probabilistic learning of similarity measures for tensor {PCA}, Pattern Recognition Letters 33 (10) (2012) 1364 – 1372. doi:http://doi.org/10.1016/j.patrec.2012.03.019.
URL http://www.sciencedirect.com/science/article/pii/S0167865512001018

[26] Y. Liu, Y. Liu, S. Zhong, K. C. Chan, Tensor distance based multilinear globality preserving embedding: A unified tensor based dimensionality reduction framework for image and video classification, Expert Systems with Applications 39 (12) (2012) 10500 – 10511. doi:http://doi.org/10.1016/j.eswa.2012.02.139.

Journal Pre-proof

URL http://www.sciencedirect.com/science/article/pii/S0957417412004022

[27] M. Safayani, M. T. M. Shalmani, Three-dimensional modular discriminant analysis (3dmda): A new feature extraction approach for face recognition, Computers & Electrical Engineering 37 (5) (2011) 811–823.

[28] A. Ouamane, A. Chouchane, E. Boutellaa, M. Belahcene, S. Bourennane, A. Hadid, Efficient tensor-based 2d+3d face verification, IEEE Transactions on Information Forensics and Security PP (99) (2017) 1–1. doi:10.1109/TIFS.2017.2718490.

[29] M. A. O. Vasilescu, D. Terzopoulos, Multilinear subspace analysis of image ensembles, in: 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., Vol. 2, 2003, pp. II–93. doi:10.1109/CVPR.2003.1211457.

[30] J. Wang, A. Barreto, L. Wang, Y. Chen, N. Rishe, J. Andrian, M. Adjouadi, Multilinear principal component analysis for face recognition with fewer features, Neurocomputing 73 (10) (2010) 1550 – 1555, subspace Learning / Selected papers from the European Symposium on Time Series Prediction. doi:https://doi.org/10.1016/j.neucom.2009.08.022.
URL http://www.sciencedirect.com/science/article/pii/S092523121000113X

[31] D. Tao, X. Li, X. Wu, S. J. Maybank, General tensor discriminant analysis and gabor features for gait recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (10) (2007) 1700–1715. doi:10.1109/TPAMI.2007.1096.

[32] S. Brahnam, L. C. Jain, L. Nanni, A. Lumini, Local Binary Patterns: New Variants and Applications, Vol. 506, 2014.

[33] C. H. Chan, M. A. Tahir, J. Kittler, M. Pietikinen, Multiscale local phase quantization for robust component-based face recognition using kernel fu-

Journal Pre-proof

sion of multiple descriptors, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (5) (2013) 1164–1177. doi:10.1109/TPAMI.2012.199.

[34] C. H. Chan, J. Kittler, N. Poh, T. Ahonen, M. Pietikinen, (multiscale) local phase quantisation histogram discriminant analysis with score normalisation for robust face recognition, in: 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 2009, pp. 633–640. doi:10.1109/ICCVW.2009.5457642.

[35] A. Hyvrinen, J. Hurri, P. O. Hoyer, Natural Image Statistics: a probabilistic approach to early computational vision, Vol. Vol. 39, 2009.

[36] B. Moghaddam, T. Jebara, A. Pentland, Bayesian face recognition, Pattern Recognition 33 (11) (2000) 1771 – 1782. doi:https://doi.org/10.1016/S0031-3203(99)00179-X.
URL http://www.sciencedirect.com/science/article/pii/S003132039900179X

[37] M. Kstinger, M. Hirzer, P. Wohlhart, P. M. Roth, H. Bischof, Large scale metric learning from equivalence constraints, in: 2012 IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 2288–2295. doi:10.1109/CVPR.2012.6247939.

[38] H. V. Nguyen, L. Bai, Cosine Similarity Metric Learning for Face Verification, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011, pp. 709–720. doi:10.1007/978-3-642-19309-5_55.
URL http://dx.doi.org/10.1007/978-3-642-19309-5_55

[39] C. Liu, Discriminant analysis and similarity measure, Pattern Recognition 47 (1) (2014) 359 – 367. doi:https://doi.org/10.1016/j.patcog.2013.06.023.
URL http://www.sciencedirect.com/science/article/pii/S0031320313002756

Journal Pre-proof

[40] R. Fang, K. D. Tang, N. Snavely, T. Chen, Towards computational models of kinship verification, in: 2010 IEEE International Conference on Image Processing, 2010, pp. 1577–1580. doi:10.1109/ICIP.2010.5652590.

[41] S. Xia, M. Shao, J. Luo, Y. Fu, Understanding kin relationships in a photo, IEEE Transactions on Multimedia 14 (4) (2012) 1046–1056. doi:10.1109/TMM.2012.2187436.

[42] J. P. Robinson, M. Shao, Y. Wu, H. Liu, T. Gillis, Y. Fu, Visual kinship recognition of families in the wild, in: IEEE Transactions on pattern analysis and machine intelligence, 2018.

[43] H. Yan, Kinship verification using neighborhood repulsed correlation metric learning, Image Vision Comput. 60 (C) (2017) 91–97. doi:10.1016/j.imavis.2016.08.009.
URL https://doi.org/10.1016/j.imavis.2016.08.009

[44] K. Zhang, Y. Huang, C. Song, H. Wu, L. Wang, Kinship verification with deep convolutional neural networks, in: BMVC, 2015.

[45] F. E. Harrell Jr, Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis, Springer, 2015.

[46] O. Laiadi, A. Ouamane, A. Benakcha, A. Taleb-Ahmed, A. Hadid, Learning multi-view deep and shallow features through new discriminative subspace for bi-subject and tri-subject kinship verification, Applied Intelligencedoi:10.1007/s10489-019-01489-2.
URL https://doi.org/10.1007/s10489-019-01489-2

[47] D. X. Meina Kan, Shiguang Shan, X. Chen, Side-information based linear discriminant analysis for face recognition, in: Proc. BMVC, 2011, pp. 125.1–125.0, http://dx.doi.org/10.5244/C.25.125.

[48] M. Bessaoudi, A. Ouamane, M. Belahcene, A. Chouchane, E. Boutellaa, S. Bourennane, Multilinear side-information based discriminant analysis

Journal Pre-proof

for face and kinship verification in the wild, Neurocomputing 329 (2019) 267 – 278. doi:https://doi.org/10.1016/j.neucom.2018.09.051.
URL http://www.sciencedirect.com/science/article/pii/S092523121831124X

[49] R. Gopalan, S. Taheri, P. Turaga, R. Chellappa, A blur-robust descriptor with applications to face recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence 34 (6) (2012) 1220–1226. doi:10.1109/TPAMI.2012.15.

[50] N. Kohli, M. Vatsa, R. Singh, A. Noore, A. Majumdar, Hierarchical representation learning for kinship verification, IEEE Transactions on Image Processing 26 (1) (2017) 289–302. doi:10.1109/TIP.2016.2609811.

[51] J. Hu, J. Lu, Y. Tan, Sharable and individual multi-view metric learning, IEEE Transactions on Pattern Analysis and Machine Intelligence 40 (9) (2018) 2281–2288. doi:10.1109/TPAMI.2017.2749576.

[52] X. Qin, D. Liu, D. Wang, Heterogeneous similarity learning for more practical kinship verification, Neural Processing Letters 47 (3) (2018) 1253–1269. doi:10.1007/s11063-017-9694-3.
URL https://doi.org/10.1007/s11063-017-9694-3

[53] X. Zhou, K. Jin, M. Xu, G. Guo, Learning deep compact similarity metric for kinship verification from face images, Information Fusion 48 (2019) 84 – 94. doi:https://doi.org/10.1016/j.inffus.2018.07.011.
URL http://www.sciencedirect.com/science/article/pii/S1566253517307273

[54] A. Moujahid, F. Dornaika, A pyramid multi-level face descriptor: application to kinship verification, Multimedia Tools and Applicationsdoi:10.1007/s11042-018-6517-0.
URL https://doi.org/10.1007/s11042-018-6517-0

Journal Pre-proof

[55] J. Hu, J. Lu, Y. Tan, J. Yuan, J. Zhou, Local large-margin multi-metric learning for face and kinship verification, IEEE Transactions on Circuits and Systems for Video Technology 28 (8) (2018) 1875–1891. doi:10.1109/TCSVT.2017.2691801.

[56] S. Wang, Z. Ding, Y. Fu, Cross-generation kinship verification with sparse discriminative metric, IEEE Transactions on Pattern Analysis and Machine Intelligence (2018) 1–1doi:10.1109/TPAMI.2018.2861871.

Journal Pre-proof

Journal Pre-proof

Conflict of Interest and Authorship Conformation Form

Please check the following as appropriate:

- o All authors have participated in (a) conception and design, or analysis and interpretation of the data; (b) drafting the article or revising it critically for important intellectual content; and (c) approval of the final version.
- o This manuscript has not been submitted to, nor is under review at, another journal or other publishing venue.
- o The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript
- o The following authors have affiliations with organizations with direct or indirect financial interest in the subject matter discussed in the manuscript:

Author's name	Affiliation

Journal Pre-proof

Journal Pre-proof

Tensor Cross-view Quadratic Discriminant Analysis for Kinship Verification in the Wild

Oualid Laiadi, Abdelmalik Ouamane, Abdelhamid Benakcha, Abdelmalik Taleb-Ahmed, Abdenour Hadid

Author Biographies

Oualid Laiadi:

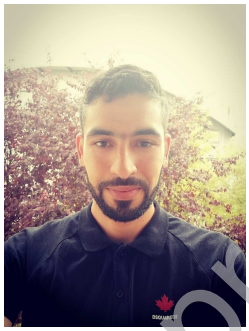


Figure 1: Oualid Laiadi.

Received the B.S. and M.S. degrees in signals and communications in 2013 and 2015, respectively, from Mohammed Khider University, Biskra, Algeria, where he is currently working toward the Ph.D. degree in the Department of Electrical Engineering. His research interests include pattern recognition, image processing, machine learning, and computer vision.

Abdelmalik Ouamane:

Figure 2: Abdelmalik Ouamane.

He received the Ph.D. degree from University of Mohammed Khider, Algeria in 2015. Currently, he is a Lecturer at Electrical Engineering Department at of the same university. He has been a reviewer for many conferences and journals. His research interests include image processing, classification, multimodal biometric and tensor analysis.

Journal Pre-proof

Journal Pre-proof

Abdelhamid Benakcha:

Figure 3: Abdelhamid Benakcha.

He had achieved a M.Sc. in 1980 and a B.Sc. in 1983 from the University of Batna, Algeria, and a Master of Science in 1985 and a Ph.D. in electronics from the University of Clermont-Ferrand, France. Since 1991, he teaches at the University of Biskra, Algeria. He is member of the Research Laboratory of electro-technique of Biskra and the head of the research group: Simulation of sliding mode control of machines. His other research interests are: Electric machines (design, modelling, identification, control), power electronics, electromagnetism (antennas, free propagation) and biometrics.

Abdelmalik Taleb-Ahmed:

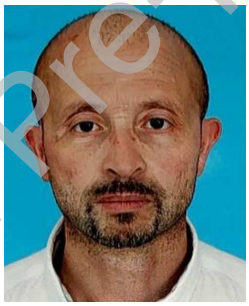


Figure 4: Abdelmalik Taleb-Ahmed.

He received the Maitrise and DEA degrees in electronics in 1987 and 1989, respectively, and the Doctorate degree in signal processing and image processing from the Université des Sciences et Technologie de Lille, France, in 1992. From 1992 to 2004, he was a Professor at the Université des Sciences et Technologie de Lille. He is currently a full Professor with IEMN, Polytechnic University of Hauts-de-France. His research interests include digital signal processing and biomedical image processing.

Journal Pre-proof

Journal Pre-proof

Abdenour Hadid:

Figure 5: Abdenour Hadid.

Received the D.Sc. degree in electrical and information engineering from the University of Oulu, Finland, in 2005. He is currently an Academy Research Fellow at the Center for Machine Vision and Signal Analysis, University of Oulu. He is regularly visiting the School of Electronics and Information, Northwestern Polytechnical University, Xian, China. His research interests include computer vision, machine learning, and pattern recognition with a particular focus on biometrics.

Author Affiliation'

(a) (b) (c)

Figure 6: Author affiliation', (a) University of Biskra, Biskra, Algeria; (b) Polytechnic University of Hauts-de-France, Valenciennes, France; (c) University of Oulu, Oulu, Finland.

Journal Pre-proof