Computer Vision and Image Understanding xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

Kinship verification using multiview hybrid distance learning

Shahar Mahpod*, Yosi Keller

Faculty of Engineering, Bar Ilan University, Israel

ABSTRACT

The estimation of kin relationships between parents and their children based on their face images is a common biometric task, conducted daily by human observers. Kin similarity is subject to significant appearance variability, as parents and their children differ by age and gender. In this work we propose a multiview hybrid combined symmetric and asymmetric distance learning network for facial kinship verification. Dual discriminative representations are learnt for the parents and the children using a margin maximization learning scheme, while the kin verification is formulated as a classification problem solved by SVM. The proposed scheme was successfully applied to the KinFaceW and KinFaceCornell datasets, comparing favorably with contemporary state-of-the-art approaches.

1. Introduction

The analysis of face images is a classical research topic in computer vision that was studied in a gamut of research problems, such as face recognition (Lu et al., 2013), face verification (Hu et al., 2014a; Simonyan et al., 2013), facial expression analysis (Fang et al., 2011), gender (Sun et al., 2006) and age estimation (Yan et al., 2014a), to name a few. In this work we study the kin verification problem, where given a pair of face images $\{\phi_i, \varphi_i\}$ we aim to determine whether ϕ_i and φ_i are kins based on their appearance, such that $\phi_i \in \{Mother, Father\}$ and $\varphi_i \in \{Daughter, Son\}$. The problem is exemplified in Fig. 1.

Kin verification schemes typically utilize features-based face representations such that $\hat{\boldsymbol{\phi}}_i = h(\boldsymbol{\phi}_i)$ and $\hat{\boldsymbol{\varphi}}_i = h(\boldsymbol{\varphi}_i)$, where $h(\cdot)$ is an image feature that might encode eye and skin color (Fang et al., 2010), histograms of gradients (Fang et al., 2010), Weber normalization (Kohli et al., 2012), spatial pyramids (Zhou et al., 2011), Gabor-based Gradient Orientation Pyramid (GGOP) (Zhou et al., 2012), Local Binary Patterns (LBP) and Three-Patch LBP (TPLBP) (Lu et al., 2012).

A common approach to kin verification is to compute the similarity between the faces $\{\hat{\boldsymbol{\phi}}_i, \hat{\boldsymbol{\varphi}}_i\}$, by first applying Distance Learning (DL) (Lu et al., 2014; Guillaumin et al., 2009; Lu et al., 2015; Qin et al., 2015a; Zheng et al., 2015) to the image representation $\{\hat{\boldsymbol{\phi}}_i, \hat{\boldsymbol{\varphi}}_i\} \in \mathbb{R}^D$, aiming to increase the classification margin between related and unrelated pairs of image faces. It is common to learn a symmetric Mahalanobis distance

$$d_i^s = \|\mathbf{W}_S\hat{\boldsymbol{\phi}}_i - \mathbf{W}_S\hat{\boldsymbol{\varphi}}_i\|, \tag{1}$$

where $\mathbf{W}_S \in \mathbb{R}^{d\times D}$, $d \ll D$ is a matrix that projects a face descriptor $\hat{\boldsymbol{\phi}}_i$ into a low-dimensional space. Multiview DL is applied when multiple image representations are used (Cui et al., 2013; Hu et al., 2014b; Lu et al., 2012)

$$h(\boldsymbol{\phi}_i) = [h_1(\boldsymbol{\phi}_i), \ldots, h_K(\boldsymbol{\phi}_i)]^T. \tag{2}$$

Symmetric distance learning formulations such as in Eq. 1 are frequently used in face recognition and verification as in Fig. 2, that depicts representative images taken from the LFW dataset (Huang et al., 2007), where the faces are of the same person taken at relatively similar ages.

In contrast to common face recognition/verification tasks (Taigman et al., 2014), where one aims to match visually similar faces, kin verification is an inherently *asymmetric* distance learning problem (Mignon and Jurie, 2012), where we aim to match *dissimilar* images, as depicted in Fig. 1. Kin verification image pairs differ significantly by age (father vs. son in Fig. 1a) and gender (mother vs. son in Fig. 1b).

Thus, as both symmetric (Mignon and Jurie, 2012) and asymmetric DL formulations (Lu et al., 2015; Qin et al., 2015a; Zheng et al., 2015) had been shown to be efficient in kin verification, we propose a hybrid distance learning (HDL) network for kin verification, that utilizes both symmetric and asymmetric DL terms

$$d_i^h = \|\mu(\mathbf{W}_\phi\hat{\boldsymbol{\phi}}_i - \mathbf{W}_\varphi\hat{\boldsymbol{\varphi}}_i) + (1-\mu)\mathbf{W}_S(\hat{\boldsymbol{\phi}}_i - \hat{\boldsymbol{\varphi}}_i)\|_2, \tag{3}$$

$\{\mathbf{W}_\phi, \mathbf{W}_\varphi, \mathbf{W}_S\} \in \mathbb{R}^{d\times D}$ are the learnt projection matrices, and $\mu \in [0, 1]$ is an autotuned weighting factor. The HDL parameters are trained by applying Stochastic Gradient Descent (SGD) to a Margin Maximization functional based on the Hinge Loss functional. It is extended to a multiview formulation using a network framework, by first applying

* Corresponding author.
E-mail addresses: mahpod.shahar@gmail.com (S. Mahpod), yosi.keller@gmail.com (Y. Keller).

https://doi.org/10.1016/j.cviu.2017.12.003
Received 1 May 2017; Received in revised form 24 September 2017; Accepted 13 December 2017

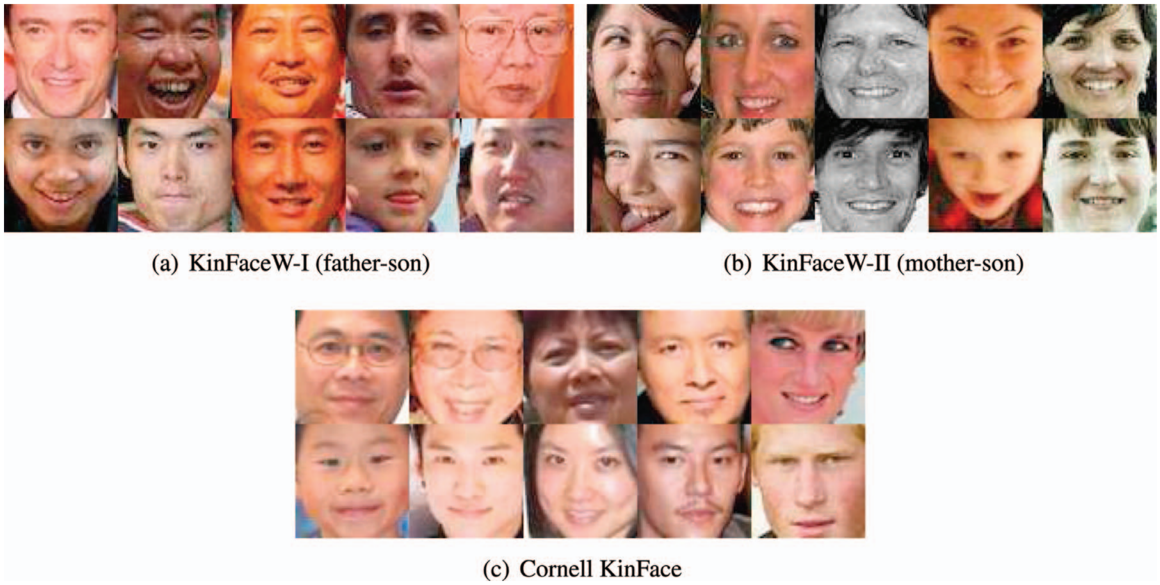

Fig. 1. Kin verification examples. The upper row depicts the parent, while the lower one shows their children. Kin verification is characterized by matching significantly dissimilar images.

Fig. 2. Facial image verification examples from the LFW dataset (Huang et al., 2007). The images are characterized by similar appearances.

the HDL formulation to multiple image cues $\left\{\hat{\boldsymbol{\phi}}_i^k, \hat{\boldsymbol{\varphi}}_i^k\right\}_1^K$ *separately*, and then fusing the results using a second HDL layer. We denote the resulting approach multiview hybrid distance learning (MHDL), where the kinship verification is given by a Kernel SVM classifier.

In that we propose the following contributions:

First, we derive a novel hybrid distance learning (HDL) scheme based on margin maximization, that utilizes both symmetric and asymmetric DL formulations and is applied to the kin verification problem.

Second, we propose a multiview kin verification network consisting of two HDL layers, denoted as Multiview Hybrid Distance Learning (MHDL), and SVM-based classification. The first HDL layer learns a HDL-based image representation per feature, while the second, computes a unified HDL-based representation of all features.

Last, the proposed kin verification scheme is experimentally shown to be robust to the choice of training parameters, and compares favorably with contemporary state-of-the-art approaches, in terms of kin-verification accuracy, when applied to the KinFaceW-I, KinFaceW-II (Lu et al., 2012), and Cornell KinFace (Fang et al., 2010) datasets.

The rest of this paper is organized as follows: we start by reviewing previous results in kin verification in Section 2. The proposed HDL and MHDL approaches are introduced in Section 3, while the experimental validation and comparison with contemporary state-of-the-art schemes is presented in Section 4. Concluding remarks and future work are discussed in Section 5.

2. Related work

Kinship verification is a variation of the face verification problem, that aims to match the faces of parents and their children. As such, it is often formulated as a binary classification task, where it is common to apply a binary classification scheme to low-level face image representations. The acquisition of publicly available kin face image datasets was instrumental in the derivation of kinship estimation schemes. Xia et al. (2011, 2012) introduced the UB KinFace Ver2.0 dataset, consisting of 600 images of parents at varying ages, and their children. They propose to utilize the similarity between the image faces of young parents and their children, and relate it to the similarity between the elder parents and their face images at a younger age. The *SiblingsDB* kin face database was acquired and made public by Vieira et al. (2014). It consists of high quality images of sibling pairs that were taken under controlled conditions. Guo et al. (2014) introduced a database of group photographs with corresponding kinship annotations collected from the KinFaceW-II and children-Face datasets. They proposed to agglomerate the facial similarities between all family members in a photograph, to improve the verification accuracy using a graph-based approach. The *Family101* dataset was acquired by Dehghan et al. (2014), and neural network auto-encoders were applied to automatically detect facial features and determine the kin relationships. The kinship relationship was estimated using a single image of a child, and a group of images of the candidate parent, using a SVM classifier. The *Family101* dataset was also studied by Dong et al. (2014) who proposed to extract facial patches, HOG, SIFT and FPLBP features, and apply SVM for kinship classification. A large scale kinship image database, denoted as Families in the Wild (FIW) was collected by Robinson et al. (2016), and will be made publicly available in the future.

The “FG 2015 Kinship Verification” challenge organized by Lu et al. (2015) provided a review of contemporary kin verification schemes, alongside the KinFaceW dataset and a detailed test protocol that is used in the experimental verification in Section 4. Lpez et al. (2016) noted that as most of the image kinship pairs in that dataset were cropped from the same photographs, using a simple feature such as Lab color space allows to achieve kinship verification accuracy comparable to state-of-the-art methods.