Pattern Recognition 000 (2017) 1–10

Contents lists available at ScienceDirect

Pattern Recognition

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

Video-based kinship verification using distance metric learning

Haibin Yan[a,*], Junlin Hu[b]

[a] School of Automation, Beijing University of Posts and Telecommunications, Beijing, 100876, China
[b] School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 639798, Singapore

ARTICLE INFO

Article history:
Received 1 September 2016
Revised 7 January 2017
Accepted 1 March 2017
Available online xxx

Keywords:
Kinship verification
Metric learning
Face recognition
Video-based

ABSTRACT

In this paper, we investigate the problem of video-based kinship verification via human face analysis. While several attempts have been made on facial kinship verification from still images, to our knowledge, the problem of video-based kinship verification has not been formally addressed in the literature. In this paper, we make the two contributions to video-based kinship verification. On one hand, we present a new video face dataset called Kinship Face Videos in the Wild (KFVW) which were captured in wild conditions for the video-based kinship verification study, as well as the standard benchmark. On the other hand, we employ our benchmark to evaluate and compare the performance of several state-of-the-art metric learning based kinship verification methods. Experimental results are presented to demonstrate the efficacy of our proposed dataset and the effectiveness of existing metric learning methods for video-based kinship verification. Lastly, we also evaluate human ability on kinship verification from facial videos and experimental results show that metric learning based computational methods are not as good as that of human observers.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Kinship verification from human faces in a relatively new problem in biometrics in recent years. The key motivation of this research topic is from the research observations in results in psychology and cognitive sciences [1–4] where human faces convey an important cue for kin similarity measure because children usually look like their parents. Verifying human kinship relationship has several potential applications such as image annotation, family album organization, social media mining, and missing children searching. Over the past few years, a number of kinship verification methods have been proposed in the literature, which aims to present effective computational models to verify human kinship relations via facial image analysis [5–17]. While these methods have achieved some encouraging performance [5–13,15,18], it is still challenging to develop discriminative and robust kinship verification approached for real-world applications, especially when face images are captured in unconstrained environments where large variations of pose, illumination, expression, and background occurs.

Most existing kinship verification methods determine human kinship relationship from still face images. Due to the large variations of human faces, a single still image may not be discriminative enough to verify human kin relationship. Compared to a single image, a face video provides more information to describe the appearance of human face. It can capture the face of the person of interest from different poses, expressions, and illuminations. Moreover, face videos can be much easier captured in real applications because there are extensive surveillance cameras installed in public areas. Hence, it is desirable to employ face videos to determine the kin relations of persons. However, it is also challenging to exploit discriminative information of face videos because intra-class variations are usually larger within a face video than a single sill image.

In this paper, we investigate the problem of video-based kinship verification via human face analysis. Specifically, we make the two contributions to video-based kinship verification. On one hand, we present a new video face dataset called Kinship Face Videos in the Wild (KFVW) which were captured in wild conditions for the video-based kinship verification study, as well as the standard benchmark. On the other hand, we employ our benchmark to evaluate and compare the performance of several state-of-the-art metric learning based kinship verification methods. Experimental results are presented to demonstrate the efficacy of our proposed dataset and the effectiveness of existing metric learning methods for video-based kinship verification. Lastly, we also test human ability on kinship verification from facial videos and experimental results show that metric learning based computational methods are not as good as that of human observers.

* Corresponding author.
E-mail address: eyanhaibin@bupt.edu.cn (H. Yan).

http://dx.doi.org/10.1016/j.patcog.2017.03.001
0031-3203/© 2017 Elsevier Ltd. All rights reserved.

Table 1
Review and summary of existing representative kinship verification methods in the literature.

Method	Characteristics	Type	Year
Fang et al. [5]	Local feature representation	image	2010
Zhou et al. [6]	Local feature representation	image	2011
Xia et al. [7]	Transfer subspace learning	image	2012
Guo and Wang [9]	Bayes inference	image	2012
Zhou et al. [10]	Local feature representation	image	2012
Kohli et al. [18]	Local feature representation	image	2012
Somanath et al. [11]	Local feature representation	image	2012
Dibeklioglu et al. [19]	Dynamic feature representation	image	2013
Lu et al. [13]	Distance metric learning	image	2014
Guo et al. [14]	Logistic regression	image	2014
Yan et al. [15]	Multi-metric learning	image	2014
Yan et al. [20]	Mid-feature learning	image	2015
Our work	Distance metric learning	video	

The rest of this paper is organized as follows. In Section 2, we briefly review some related work, and Section 3 introduces the Kinship Face Videos in the Wild (KFVW) dataset. Section 4 presents some popular metric learning methods which have been widely used in kinship verification. Section 5 presents the experimental results and analysis. Finally, Section 6 concludes this paper.

2. Related work

In this section, we briefly review the related topics to our work: (1) kinship verification, (2) metric learning, and (3) video-based face analysis.

2.1. Kinship verification

The first study on kinship verification from facial images was made in [5]. In their work, they extracted local features such as skin color, gray value, histogram of gradient, and facial structure information in facial images and select some of them for kinship verification. Since this seminal work, more and more kinship verification methods have been proposed in the literature [5–7,9,10,13,15,18–23]. These methods can be mainly categorized into two classes: feature-based [5,6,9,10,19] and model-based [7,13–15]. Methods in the first class extract discriminative feature descriptors to represent kin-related information. Representative such feature information include skin color [5], histogram of gradient [5,6,11], Gabor wavelet [7,10,11,24], gradient orientation pyramid [10], local binary pattern [13,25], scale-invariant feature transform [11,13,15], salient part [8,9], self-similarity [18], and dynamic features combined with spatio-temporal appearance descriptor [19]. Methods in the second class learn discriminative models to verify kin relationship from face pairs. Typical such models are subspace learning [7], metric learning [13,15], transfer learning [7], multiple kernel learning [10] and graph-based fusion [14]. Table 1 lists a review and summary of existing representative kinship verification methods in the literature. All these kinship verification methods determine human kinship relationship from still face images, which may not discriminative enough to verify human kin relationship since large variations of human faces usually occur in still images.

2.2. Metric learning

A variety of metric learning methods [26–28,28–52] have been widely used in numerous computer vision tasks such as face recognition [26,28], gait recognition [34], object recognition, human activity recognition [29], human age estimation [30], person re-identification [28,31,32], visual tracking, and visual search. These methods can be mainly classified into two classes: unsupervised and supervised. The first class of methods learn a low-dimensional manifold to preserve the geometrical structure of data points, and the second class of methods seek an appropriate distance metric to exploit the discriminative information of samples. Recently, metric learning techniques have also been used in kinship verification [13,15], these methods are strongly supervised and require the exact label information of samples. For kinship verification, it is more convenient to obtain the weakly supervision of samples so that it is desirable to employ and evaluate weakly supervised methods for kinship verification.

Table 2
Compassion of existing facial datasets for kinship verification.

Dataset	Number of kinship pairs	Type	Year
CornellKin [5]	150	image	2010
UB KinFace [7]	400	image	2012
IIITD Kinship [18]	272	image	2012
Family101 [12]	206	image	2013
KinFaceW-I [13]	533	image	2014
KinFaceW-I [13]	1000	image	2014
TSKinFace [66]	2030	image	2015
KFVW (Ours)	418	video	

2.3. Video-based face recognition

A variety of video-based face analysis methods have been proposed in the literature, and these methods can be mainly classified into parametric [53–56] and nonparametric [57–65] methods. Parametric methods represent each face video as a parametric family of probabilistic distribution, and use the Kullback–Leibler divergence to measure the similarity of two face videos. However, these methods usually fail when the underlying distributional assumptions do not hold for different face videos. Nonparametric methods exploit geometrical information to measure the similarity of two face videos by modeling face each video as a single linear subspace or as the union of linear subspaces. While a variety of video-based face recognition methods have been presented, there is no work on video-based kinship verification, probably due to the lack of such datasets. In this work, we fill this gap and contribute a video dataset for kinship verification.

3. The kinship face videos in the wild dataset

In past few years, several facial datasets have been released to advance the kinship verification problem, e.g., CornellKin [5], UB KinFace [7], IIITD Kinship [18], Family101 [12], KinFaceW-I [13], KinFaceW-II [13], TSKinFace [66], etc. Table 2 provides a summary of existing facial datasets for kinship verification. However, these datasets only consist of still face images, in which each subject usually has a single face image. Due to the large variations of human faces, a single still image may not be discriminative enough to verify human kin relationship. To address these shortcomings, we collected a new video face dataset called Kinship Face Videos in the Wild (KFVW) for the video-based kinship verification study. Compared to a still image, a face video provides more information to describe the appearance of human face, because it can easily capture the face of the person of interest from different poses, expressions, and illuminations.

The KFVW dataset was collected from TV shows on the Web. We totally collected 418 pairs of face videos, and each video contains about 100 – 500 frames with large variations such as pose, lighting, background, occlusion, expression, makeup, age, etc. The average size of a video frame is about 900 × 500 pixels. There are four kinship relation types in the KFVW dataset: Father-Son (F-S), Father-Daughter (F-D), Mother-Son (M-S), and Mother-Daughter (M-D), and there are 107, 101, 100, and 110 pairs of kinship face videos for kin relationships F-S, F-D, M-S, and M-D respectively.

Please cite this article as: H. Yan, J. Hu, Video-based kinship verification using distance metric learning, Pattern Recognition (2017), http://dx.doi.org/10.1016/j.patcog.2017.03.001

Download English Version:

https://daneshyari.com/en/article/6939616

Download Persian Version:

https://daneshyari.com/article/6939616

Daneshyari.com