Image and Vision Computing 32 (2014) 761–770

Contents lists available at ScienceDirect

Image and Vision Computing

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

A framework for joint estimation of age, gender and ethnicity on a large database ☆

Guodong Guo [a,*], Guowang Mu [b]

[a] Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, WV 26506, USA
[b] School of Science, Hebei University of Technology, Tianjin 300401, PR China

ARTICLE INFO

Article history:
Received 31 May 2013
Received in revised form 10 April 2014
Accepted 30 April 2014
Available online 10 May 2014

Keywords:
Joint estimation of age, gender and ethnicity
A general framework
Partial least squares (PLS)
Canonical correlation analysis (CCA)
Regularized CCA
Kernel PLS (KPLS)
Kernel CCA (KCCA)
Regularized KCCA
Least squares CCA

ABSTRACT

Human age, gender and ethnicity are valuable demographic characteristics. They are also important soft biometric traits useful for human identification or verification. We present a framework that can estimate the three traits jointly. The joint estimation framework could deal with the mutual influence of age, gender, and ethnicity implicitly. Under this joint estimation framework, we explore different methods for simultaneous estimation of age, gender, and ethnicity. The canonical correlation analysis (CCA) based methods, and partial least squares (PLS) models are explored under our joint estimation framework. Both the linear and nonlinear methods are investigated to measure the performance. We also validate some extensions of these methods, such as the least squares formulations of the CCA methods. We found some consistent ranking of these methods under our joint estimation framework. More importantly, we found that the CCA based methods can derive an extremely low dimensionality in estimating age, gender and ethnicity. An analysis of this property is given based on the rank theory. The experiments are conducted on a very large database containing more than 55,000 face images.

1. Introduction

Recently human age estimation in face images has become an active research in computer vision and pattern recognition [1,2], because of many potential applications in the real world. Age estimation is useful for creating an age-specific human–computer interaction (AS-HCI) system [3], electronic customer relationship management (ECRM) [4], and business intelligence [5].

In addition to age estimation, face images can also be used to extract gender and ethnicity information. The three major characteristics, i.e., age, gender and ethnicity, are valuable demographic information of an individual or statistics about a population.

Automated estimation of the demographics is of great value in practice, such as business intelligence, local community planning, and new school locating [6]. Age, gender and ethnicity are also useful soft biometric traits that can be used for human identification or verification.

However, current computational techniques are still not robust enough for practical uses. Thus it is demanding to develop a robust and effective system to recognize age, gender and ethnicity for a given individual or a population.

In the literature, there are different methods for the estimation of each single trait, e.g., age estimation [7–12,1], gender classification [13–17], or ethnicity estimation [18–21]. However, there are very few approaches to estimate all three traits together.

1.1. Different frameworks

In Ref. [22], a classification of ethnicity, gender, and age groups was executed, with each trait classified independently. This kind of approach can be illustrated by the framework shown in Fig. 1. An implicit assumption is that the three traits can be classified independently, and there is no relation among the three traits.

In our previous studies [23,24], we found that age estimation can be influenced by the gender and ethnicity differences significantly. In Ref. [24], we show that the age estimation errors can be increased significantly on the Yamaha age and gender database, when the males and females are mixed together for age estimation. This result explains why the previous approaches, e.g., Refs. [9–12], executed the age estimation for males and females, separately, on the Yamaha database. In Ref. [23], we studied the influence of gender and ethnicity on age estimation systematically, using the MORPH database [25]. We found

☆ This paper has been recommended for acceptance by Stan Sclaroff.
* Corresponding author.
E-mail address: guodong.guo@mail.wvu.edu (G. Guo).

http://dx.doi.org/10.1016/j.imavis.2014.04.011
0262-8856/

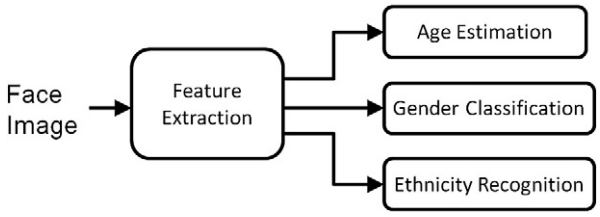

Fig. 1. A framework of independent estimation of age, gender, and ethnicity, separately.

that the age estimation errors can be increased when the estimation is performed across gender, ethnicity, or across both.

To deal with the influence of gender and ethnicity on age estimation, we proposed a framework [23] which has a three-step procedure on the extracted features from face images: (1) dimensionality reduction, (2) gender and ethnicity group classification, and (3) age estimation performed on each classified group. This framework can be illustrated by Fig. 2. Although this framework mainly focuses on improving age estimation, it also gets the gender and ethnicity characteristics in the second step.

In this paper, we present a new framework which can estimate the age, gender, and ethnicity jointly in a single step. This single-step framework is much simpler than our three-step procedure [23], and can deal with the influence of gender and ethnicity on age estimation implicitly. Because the labels of gender and ethnicity for each aging pattern are integrated into the single-step age estimation process, we call it *implicitly* dealing with the influence of gender and ethnicity on age estimation. In contrast, an *explicit* approach is the three-step procedure [23] where the gender and ethnicity groups are recognized first, before performing age estimation. The new framework of single-step, joint estimation has not been investigated before by other researchers, to the best of our knowledge.

1.2. Exploring different methods under the joint estimation framework

We explore several different methods under our new framework for a single-step, joint estimation of age, gender, and ethnicity. The basic idea of our approaches is to explore the multi-label regression formulation for the joint estimation problem.

Based on the multi-label regression formulation, our joint estimation framework is very simple, as shown in Fig. 3.

To solve the multi-label regression problem, we explore two broad categories of methods, i.e., the partial least squares (PLS) models [26,27] and canonical correlation analysis (CCA) based methods [28,29]. We found that these methods can do dimensionality reduction and joint estimation of age, gender, and ethnicity all together within a single step. This is an interesting observation in our problem, and may inspire new explorations of these methods for other pattern recognition problems.

Specifically, we found that the canonical correlation analysis (CCA) based methods, including linear CCA, regularized CCA, and kernel CCA, can find only three basis vectors to project the original features of several thousand dimensions. Thus, only three dimensions of features are needed (after the transformation) to estimate all three traits. This is a *novel* finding in estimating age, gender and ethnicity [6]. Further, we use the rank theory to analyze the feature dimensionality problem in using the CCA based methods for our specific task. Hopefully, our analysis may inspire more investigations for other recognition problems to derive similar results with a minimum number of feature dimensions.

The PLS based methods, i.e., linear PLS and kernel PLS, can reduce the dimensionality to 20 or 30 from thousands (in an original feature space), and can achieve a good performance to estimate all three traits [30]. However, the dimensionality based on PLS methods cannot reach a small number as the CCA based methods.

Under our joint estimation framework, we compare the CCA based methods with the PLS based, and derive a ranking of these methods in solving the joint estimation problem.

Further, we explore some extensions of the CCA based methods, such as the recent development of least square formulations [31]. The least square CCA methods have shown excellent performance on traditional machine learning databases, however, it is unknown whether these methods can perform well in our problem.

Given face images, a high dimensionality could be resulted in, e.g., thousands of features are extracted using the biologically-inspired features (BIFs) proposed in Ref. [12] for age estimation. Here, we show that the BIF can be used to represent the face images for all of the three characteristics, i.e., the age, gender, and ethnicity.

We also investigate whether the performance has any changes when other learning methods are used for age estimation. For this purpose, the CCA or PLS based methods are used to generate new features that have low dimensionality and discriminative capability, but not predict the ages. The aging functions are learned by other methods for age estimation.

1.3. Contributions

Our major contributions in this paper include:

1. A framework is presented for joint estimation of age, gender and ethnicity in a single step;
2. A novel finding on feature dimensionality in estimating age, gender and ethnicity;
3. A rank theory based analysis of the dimensionality problem in using the CCA based methods;

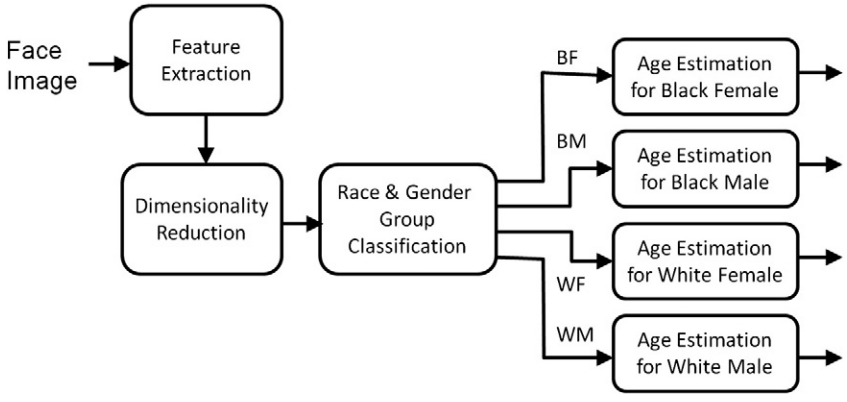

Fig. 2. The framework using a three-step procedure for age, gender, and ethnicity estimation in our previous approach [23]. After feature extraction, the three steps are dimensionality reduction, race and gender group classification, and finally age estimation within each group.