Information Sciences 441 (2018) 187–197

Contents lists available at ScienceDirect

Information Sciences

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

Face recognition based on Volterra kernels direct discriminant analysis and effective feature classification

Guang Feng [a,b], Hengjian Li [a,b,*], Jiwen Dong [a,b], Jiashu Zhang [c]

[a] School of Information Science and Engineering, University of Jinan, Jinan 250022, China
[b] Shandong Provincial Key Laboratory of Network based Intelligent Computing, Jinan 250022, China
[c] Sichuan Key Laboratory of Signal and Information Processing, Southwest Jiaotong University, Chengdu 610031, China

A R T I C L E   I N F O

Article history:
Received 18 September 2016
Revised 29 January 2018
Accepted 15 February 2018
Available online 15 February 2018

Keywords:
Face recognition
Feature extraction
Volterra kernels
Direct discriminant analysis
Voting strategy

A B S T R A C T

We present a novel face recognition method based on direct discriminant Volterra kernels and effective feature classification (DD-VK). One of the crucial steps involves dividing face images into patches and using the DD-VK method to extract the features of sub-image patches. DD-VK implements diagonalization to discard useless information in the null space of the inter-class scatter matrix and preserve important discriminant information in the null space of the intra-class scatter matrix. This method can simultaneously maximize inter-class distances and minimize intra-class distances in the feature space. We also introduce a novel classification scheme associated with the 2D Volterra kernel feature. Our scheme aggregates the classification information obtained from each column of the feature matrix in each image patch and uses a voting strategy to implement parent face image classification. This procedure can reduce the influence of local negative information. Experimental results show that the proposed method demonstrates good performance when dealing with conventional face recognition problems and exhibits strong robustness when dealing with block occlusion images.

1. Introduction

Face recognition has been widely studied due to its applications in many fields, such as security monitoring, human–computer interaction, internetwork communication, computer entertainment, artificial intelligence, and e-commerce security. Although many scholars have achieved significant advances in face recognition, the technology still encounters numerous challenges. These challenges include intra-class variations in illumination, pose, expression, noise, and occlusion. Therefore, robust and discriminant feature generation has become a problem in face recognition.

Many face feature representation methods have been proposed by scholars, and these include subspace-based global and local features [18]. Conventional global feature representation methods mainly include principal component analysis (PCA) [36], linear discriminant analysis (LDA) [5], and locality preserving projection (LPP) [12]. PCA uses orthogonal transformation to project data in a high-dimensional space into a low-dimensional space to obtain the main features of image data. LDA also projects data from a high-dimensional space to a low-dimensional one. To implement the projection, the method seeks for a set of optimal discriminant projection vectors that contains the maximum between-class scatter and the minimum within-class scatter among all projection spaces. LPP can be regarded as the linearization of Laplacian EigenMap [4]. It considers

* Corresponding author.
E-mail address: ise_lihj@ujn.edu.cn (H. Li).

https://doi.org/10.1016/j.ins.2018.02.028

the manifold structure of data and saves the location of data in a low-dimensional feature space. However, these methods demonstrate reduced efficiency when dealing with expression and illumination changes. Several advanced methods based on global feature representation [3,6,16,17,19,25,44] were proposed recently to enhance the discriminative capability of the learned feature.

Local features are more stable than global features when they are used to deal with local changes, such as illumination, expression, and inaccurate alignment. Typical methods involving local features include local binary patterns (LBP) [1,24], Gabor features [23,46], and their combination [21,37]. LBP uses the gray value of the center pixel as the threshold. The binary codes obtained through a comparison with the neighborhood are utilized to describe the local texture characteristics. This method is robust to illumination changes within a certain range. Gabor wavelets capture the local structure that corresponds to a specific spatial frequency (scale), spatial locality, and selective orientation and have been demonstrated to be discriminative and robust to illumination and expression changes. Kumar et al. [14,15] proposed a face recognition method based on Volterra kernels in 2009. They used Volterra kernels to identify the best projection of face images because these kernels can systematically produce good approximations of such a projection. Their method is significantly different from fixed-feature traditional approaches (Gabor features, LBP features, etc.) because the convolution filters utilized are not fixed a priori; instead, they are learned in a data-driven manner in conjunction with the discriminator. To elaborate, Kumar et al. try to find the best filter that groups elements of the same class together and drives elements of different classes apart.

Classifier learning is another important face recognition problem. Among various classification methods, the nearest neighbor (NN) classifier is the most popular in face recognition. Moghaddam et al. [27] converted a multi-class classification problem into a two-class classification problem by constructing interpersonal and extrapersonal variations. A query image for a particular subject can be in two situations only, namely, belonging or not belonging to that subject. Many binary classification methods, such as Bayesian [27], support vector machine (SVM) [42], and Adaboost [30], have been proposed on the basis of this condition.

Sparse representation based face classification (SRC) has also produced good results [39]. SRC linearly represents a probe image by using all of the training images under the L1-norm-based sparsity constraint. Many methods have been applied to this strategy, and these methods demonstrate state-of-the-art performance in face recognition with image corruption [31,39], face disguises [32,43,45], and minimal training data [7]. However, due to the constraint of the L1-norm-based sparsity constraint, these methods are more complex than NN, SVM, and others.

We present a novel face recognition method based on direct discriminant Volterra kernels and effective feature classification (DD-VK). First, in the feature extraction stage, we obtain the projection matrix based on Volterra kernels through eigenvalue decomposition to avoid matrix inversion operations. This method can extract identification information of data to the maximum extent. Second, in the classification stage, we propose a new method to individually classify each column of the feature matrix in each image patch. A voting strategy is used to accomplish parent image classification for robust results. Experimental results show that the proposed DD-VK outperforms other common solutions to conventional face recognition problems and demonstrates good robustness against block occlusion.

The rest of this paper is organized as follows. We introduce the theory of Volterra kernels and its application in face recognition in Section 2. And then we describe the proposed direct discriminant Volterra kernels (DD-VK) method and new classification scheme associated with DD-VK in detail in Section 3. Extensive experimental results are reported in Section 4, and conclusions are drawn in Section 5.

2. Volterra kernels

The Volterra series is essentially a functional series expansion of a nonlinear time-invariant system from the mathematical point of view. It can also be considered a 1D convolution of a linear system for promotion to multidimensional convolution space [33]. This chapter describes the Volterra series and its application in face recognition fields.

2.1. Volterra series

The Volterra series can be presented as a function of input $x(n)$ by using N-sample memory pth-order truncated Volterra series expansion. The discrete time representation is written as follows:

$$y(k) = h_0 + \sum_{p=1}^{L} \sum_{k_1=0}^{N_p-1} \cdots \sum_{k_p=k_p-1}^{N_p-1} h_p(k_1, \cdots k_p) x(n-k_1) \cdots x(n-k_p), \tag{1}$$

where $x(n)$ is the input signal, $y(k)$ is the output signal, N_p is the pth-order nonlinear memory depth, h_0 is an offset constant, and $h_p(k_1, \ldots, k_p)$ is the p-th order kernel function of a discrete Volterra series. In matrix form, Eq. (1) can be written as follows:

$$y(n) = h_v^T(n) x_v(n) \tag{2}$$

and