Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

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

Learning iterative quantization binary codes for face recognition

Lei Tian, Chunxiao Fan, Yue Ming *

Beijing University of Posts and Telecommunications, Beijing Key Laboratory of Work Safety Intelligent Monitoring, School of Electronic Engineering, No. 10 Xitucheng Road, Beijing 100876, PR China

ARTICLE INFO

Article history:
Received 8 October 2015
Received in revised form
12 March 2016
Accepted 17 June 2016
Communicated by Yue Gao

Keywords:
Iterative quantization
Feature learning
Binary coding
Local descriptor
Face recognition
Face verification

ABSTRACT

Binary feature descriptors have been widely used in computer vision field due to their excellent discriminative power and strong robustness, and local binary patterns (LBP) and its variations have proven that they are effective face descriptors. However, the forms of such binary feature descriptors are predefined in the hand-crafted way, which requires strong domain knowledge to design them. In this paper, we propose a simple and efficient iterative quantization binary codes (IQBC) feature learning method to learn a discriminative binary face descriptor in the data-driven way. Firstly, similar to traditional LBP method, we extract patch-wise pixel difference vectors (PDVs) by computing and concatenating the difference between center patch and its neighboring patches. Then, inspired by multi-class spectral clustering and the orthogonal Procrustes problem, which both are widely used in image retrieval field, we learn an optimized rotation to minimize the quantization error of mapping data to the vertices of a zero-centered binary hypercube by using iterative quantization scheme. In other words, we learn a feature mapping to project these pixel difference vectors into low-dimensional binary vectors. And our IQBC can be used with unsupervised data embedding method such as principle component analysis (PCA) and supervised data embedding method such as canonical correlation analysis (CCA), namely IQBC-PCA and IQBC-CCA. Lastly, we cluster and pool these projected binary codes into a histogram-based feature that describes the co-occurrence of binary codes. And we consider the histogram-based feature as our final feature representation for each face image. We investigate the performance of our IQBC-PCA and IQBC-CCA on FERET, CAS-PEAL-R1, LFW and PaSC databases. Extensive experimental results demonstrate that our IQBC descriptor outperforms other state-of-the-art face descriptors.

1. Introduction

During the past decades, face recognition has been successfully applied in many fields, such as access control, ID authentication, watch-list surveillance and so on. It has still attracted much attention due to its theoretical and practical challenges. Generally, there are two critical problems on conventional face recognition system: *feature representation* and *classifier training*. Recently, most of the existing works are focusing on these two aspects to improve the performance of face recognition methods when they faced with a variety of intra-class variabilities. For face representation, the purpose is to extract discriminative (global/local) features to make face images of different individual more separable. In other words, the further distance between feature with different individual is, the better feature representation method will be. For classifier training, the goal is to design an efficient (supervised/unsupervised) classifier to distinguish different face patterns.

In this paper, we mainly focus on the issues of face feature representation. Feature representation is an important problem, and it significantly affects the performance of face recognition system due to the large variations caused by the expression, pose, illumination, aging and so on. These intra-class variabilities reduce the similarity of face images from the same individual, even the intra-class variabilities often larger than the inter-class variabilities in many benchmark datasets. Up to now, various face representation methods have been proposed, including subspace-based holistic features and local features. The subspace-based holistic feature methods consider a face image as a long vector and learn the statistical information of input face images. The representative examples of holistic features include principal component analysis (PCA) [1], linear discriminant analysis (LDA) [2] and independent component analysis (ICA) [3]. Local features representation methods, as opposed to holistic features, describe the pattern of each local region of a face image and combine the described information of all regions into a final feature representation. The representative examples of local features include Gabor

* Corresponding author.
E-mail addresses: tianlei189@sina.com (L. Tian), fcxg100@163.com (C. Fan), myname35875235@126.com (Y. Ming).

http://dx.doi.org/10.1016/j.neucom.2016.06.046
0925-2312/

[4] wavelets and local binary pattern (LBP) [5]. Though, these features have achieved great success for some controlled scenarios through designing low-level features elaborately, they cannot achieve excellent performance when they faced with extreme intra-class variability and uncontrolled scenarios. Since most of these face representation methods are hand-crafted, they can achieve excellent performance only when strong prior knowledge is provided. Therefore, it is a challenging problem in face recognition that how to extract robust and discriminative feature when the intra-class variabilities are large.

Learning features from data itself instead of manually designing features is considered as a plausible way to overcome the limitation of hand-crafted features. In this paper, we propose a simple and efficient iterative quantization binary codes (IQBC) to learn a discriminative binary face descriptor in the data-driven way. Inspired by the research that binary codes are robust to local variations, we intend to learn discriminative and robust binary codes from raw pixels by iterative quantization methods, which connects the multi-class spectral clustering problem with the orthogonal Procrustes problem. Firstly, we extract patch-wise pixel difference vectors (PDVs) by computing the difference between center patch and its neighboring patches and grouping them as pixel difference matrix (PDM). Then, we perform unsupervised data embedding method (such as PCA) or supervised data embedding method (such as CCA) to reduce the dimensionality of the original PDVs. Therefore, we name our unsupervised and supervised IQBC method as IQBC-PCA and IQBC-CCA, respectively. Next, in order to minimize the quantization error between original PDVs and binary codes, we use the iterative quantization scheme to refine the initial orthogonal transformation (i.e., random orthogonal transformation). In other words, our methods can learn discriminative and robust feature representations under the condition that the energy of real-valued PDVs is preserved as much as possible by the iterative quantization scheme. Lastly, we cluster these binary vectors by the conventional K-Means method to learn the dictionary (i.e., codebook) and pool them as a bin to obtain the histogram-based feature representations of input images. Fig. 1 illustrates the whole procedure of our IQBC feature learning method. Extensive experiments demonstrate that our IQBC descriptors (IQBC-PCA and IQBC-CCA) outperform most of the state-of-the-art face representation methods.

In preparation of this paper, we are aware of a concurrent works CBFD [6]. It learns unsupervised binary codes which satisfy three terms: (1) the variance of binary codes is maximized, (2) the quantization loss is minimized and (3) the binary codes evenly distribute at each learned bin. Though both our methods and CBFD learn compact binary codes for face recognition, there are two main differences between ours and CBFD. On the one hand, CBFD optimizes the above three terms simultaneously, so that the solution about the quantization loss term is not ensured to be optimized in fixed iterations. However, our methods can ensure the quantization loss is minimized by learning a optimized rotation in relatively small iterations. Therefore, the identity information of face image can be efficiently preserved as much as possible. On the other hand, CBFD is an unsupervised feature learning method. And our method can learn binary feature representation by both unsupervised and supervised manners. Compared with unsupervised method, our method can exploit the label information to better capture the semantic structure of the dataset.

Contributions: The contributions of our work are summarized as follows:

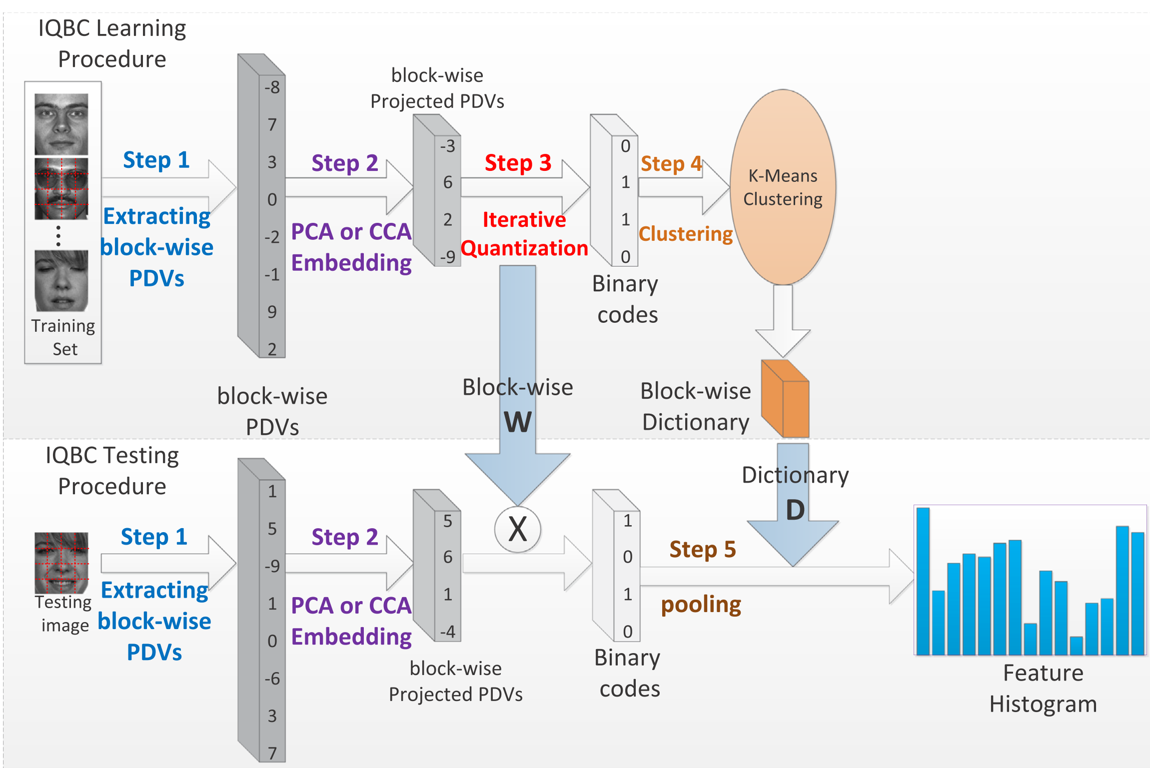

Fig. 1. The whole procedure of our proposed IQBC face representation method. For the training stage, in the first step, we extract PDVs from each local region for each training face image. In the second step, we perform unsupervised/supervised data embedding method on PDVs. In the third step, we learn a discriminative feature mapping using Iterative Quantization scheme and project dimension-reduced PDVs into low-dimensional binary codes. In the fourth step, these learned binary codes are clustered to learn a dictionary. For the testing stage, we perform the same operation as the first and the second step of training stage. Then, the PDVs are encoded into binary codes using learned mapping matrix W. Lastly, these binary codes are pooled as block-wise histograms with the learned dictionary D and these local block-wise histograms are concatenated to form the final representation of the test face image.