Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

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

Structured regularized robust coding for face recognition

Xing Wang [a], Meng Yang [a,b,*], Linlin Shen [a]

[a] College of Computer Science & Software Engineering, Shenzhen University, Shenzhen, China
[b] School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China

ARTICLE INFO

Article history:
Received 28 October 2015
Received in revised form
27 March 2016
Accepted 26 June 2016
Communicated by: Ran He

Keywords:
Structured regularized
Robust coding
Sparse representation
Pixel weight learning
Face recognition

ABSTRACT

The sparse representation based classifier (SRC) has been successfully applied to robust face recognition (FR) with various variations. To achieve much stronger robustness to facial occlusion, regularized robust coding (RRC) was proposed by designing a new robust representation residual term. Although RRC has achieved the leading performance, it ignores the structured information (i.e., spatial consistence) embedded in the occluded pixels. In this paper, we proposed a novel structured regularized robust coding (SRRC) framework, in which a weight value is assigned to each pixel to measure its importance in the coding procedure and the spatial consistence of occluded pixels is exploited by the pixel weight learning (PWL) model. Efficient algorithms were also proposed to fast learn each pixel's weight value. The experiments on face recognition in several representative datasets clearly show the advantage of the proposed SRRC in accuracy and efficiency.

1. Introduction

Face recognition (FR) is one of the hottest research topics in computer vision and pattern recognition due to its wide range of applications such as law enforcement, information security, surveillance, etc. In the past two decades FR has been extensively studied [5,23,24], and many representative methods, such as Eigenfaces [6], Fisherfaces [6], LBP [7], have been proposed. Besides, to meet the requirement of practical FR systems, active appearance model [25,26] and active shape model [27] were introduced to handle face alignment. LBP [7] and its extensions [28–31] were used to extract variation-robust texture features. To cope with the illumination variations, Cheng et al. [32] presented a robust face recognition method based on illumination invariant feature in nonsubsampled contourlet transform domain. Wu et al. [33] generalized the Weber-face to multi-scale versions for illumination-robust face recognition. Based on the facial symmetry Hsieh and Tung [34] proposed to utilize the shadow compensation to overcome illumination changes. In terms of dealing with facial occlusion, Eigenimages [8,9], probabilistic local approaches [10], Markov random fields [19] and level set method [35] were proposed for FR with occlusion. A robust kernel representation model with statistical local features was developed to handle misalignment, pose variation, and occlusion in face images [36]. In addition, statistical learning of local features [37] was proposed for FR with partial variations, such as expression and occlusion. Although much progress has been made, robust FR to occlusion/disguise still remains challenging due to the complex occlusion variations such as different categories of disguises and the unknown intensity of occluded pixels.

Inspired by studies of the role of parsimony in human vision system [39,40] and the successful application of sparse presentation in compressive sensing [41] and image restoration [42], Wright et al. [1] originally applied sparse presentation to FR and proposed the sparse representation based classifier (SRC). SRC assumes that samples of different subjects lie on different low dimensional subspaces. In SRC, a testing sample is sparsely coded over all training samples of all subjects and l_1-norm is imposed on the coding coefficients to guarantee the sparsity. Then the testing sample is classified into the class yielding the minimal reconstruction residual. In addition, SRC uses l_1-norm to characterize the coding residuals to deal with FR with occlusion. Shortly afterward, Zhang et al. [2] argued that what actually matters is not the l_1-norm regularization on the coding coefficients but the collaborative representation mechanism (i.e. using all training samples to collaboratively represent the testing sample) and proposed the collaborative representation based classifier (CRC) using l_2-norm regularization on the coding coefficients. It was shown in [2] that CRC can achieve comparable performance with SRC for FR without outlier pixels but is more efficient. The limitation of CRC is the absence of mechanisms to handle FR with occlusion.

Following SRC and CRC, many representation based works have been proposed for the general robust representation problem or the specific FR with occlusion problem. [18] developed a half-quadratic

* Corresponding author.
E-mail address: yang.meng@szu.edu.cn (M. Yang).

http://dx.doi.org/10.1016/j.neucom.2016.06.065
0925-2312/© 2016 Elsevier B.V. All rights reserved.

(HQ) framework to solve the general robust sparse representation problem. The proposed framework unifies algorithms for error correction and detection by defining different kinds of half-quadratic functions, specifically using the additive form for error correction and multiplicative form for error detection, respectively. [17] proposed a structured sparse error coding (SSEC) model to explore the intrinsic structure of the occlusion error in the face image. Specifically, it uses a morphological graph model for the spatial continuity and the shape of the occlusion error and a exponential probabilistic model for the error distribution. Du et al. [43] presented a graph regularized low-rank sparse representation recovery (GLRSRR) method to tackle the case where the facial occlusion may exist in both the training samples and the test sample. Different from SRC assuming the sparsity of the coding residuals, He et al. [11] used a Gaussian kernel-based fidelity term to regularize the coding residuals and proposed a correntropy-based sparse representation (CESR) for robust FR. And Gabor feature was also introduced in the framework of SRC to make the occlusion dictionary compressible and enhance the discrimination [12]. In order to handle more general types of facial occlusion, Yang et al. [4] designed a robust representation term and proposed a regularized robust coding (RRC) model, which has attracted much attention in the field due to its state-of-the-art performance. In RRC, each pixel is assigned a weight and outlier pixels will be assigned low weights to suppress their influence on the recognition process. [38] presented a robust nuclear norm regularized regression (RNR) model for face recognition with occlusion. RNR borrows the weight function of RRC to assign each pixel a weight and constrains the representation residual with nuclear norm to implicitly exploit the structural information of the error image. And as CRC, RNR uses the l_2-norm to regularize the coding coefficients.

Although RRC [4], CESR [11] and Gabor-SRC [12] have achieved leading performance in robust face recognition, all of them measure each pixel's representation residual independently without considering the structured information (i.e., spatial consistence) embedded in the 2D image space. In practical face recognition, most of occluded pixels are not independent but spatially consistent (e.g., block occlusion, facial disguise). Here we should note that the spatial consistence is embedded in occluded pixels but not the occluded pixels' values. Fig. 1 gives an example to show the spatial consistence of image pixels and image pixels' values.

In this paper, we use a weight to measure a pixel's importance in the coding process, then the structured information could be easily exploited in the pixel weight learning (PWL) model without considering the difference among pixels' values. With the proposed PWL model, the structured information of image pixels could be effectively exploited and a novel framework of structured regularized robust coding (SRRC) was presented for robust FR. We also present efficient algorithms to solve the PWL model. Compared with SSEC [17], HQ [18] and RNR [38], which are also structure regularized based regression methods, the proposed SRRC is very different from them. The difference between SRRC and SSEC is the type of the explored information of the occlusion and the way to explore it. SRRC exploits only the spatial continuity of the occlusion by introducing the PWL model. Whereas both the shape and the spatial continuity information of occlusion are considered in SSEC. One limitation of SSEC is that it does not always converge to the optimal solution. So an additional quality assessment model is used to select the optimal solution from the iteration sequence. The difference between SRRC and HQ lies in the specific research focus. SRRC tries to deal with robust FR with occlusion. By contrast, the purpose of HQ is to introduce a more general robust sparse model but ignores the spatial relationship between different image pixels. The difference between SRRC and RNR is the way to exploit the structural information of the occlusion. RNR imposes the nuclear norm constraint on the representation residual to make use of the structural information implicitly, whereas SRRC is more explicit and straightforward. SRRC directly requires the neighboring pixels to have similar weight values through the PWL model to ensure the local consistency of occluded pixels. Furthermore, there is one common difference between SRRC and the other three works, which is the investigation of the type of the norm imposed on the representation coefficients. In this paper we investigate the influence of different regularizations including both l_1-norm and l_2-norm on the representation coefficients and conclude that l_2-norm regularization can achieve comparable performance with l_1-norm regularization but is much faster. This conclusion is consistent with the previous works, such as CRC [2] and RRC [4]. However, SSEC and HQ only involve l_1-norm regularization and RNR only involves l_2-norm regularization. We evaluate the effectiveness of SRRC on several benchmark datasets, such as Extended Yale B [20], CMU Multi-PIE [21], AR [13] and a joint face database of AR [13] and CAS-PEAL [14]. The experiments on these datasets clearly show the advantage of SRRC in both the recognition accuracy and the computational efficiency.

The rest of this paper is organized as follows. Section 2 briefly reviews the related works. Section 3 presents the proposed structured regularized robust coding framework. Section 4 conducts the experiments, and Section 5 concludes the paper.

2. Related works

2.1. Sparse representation based classification (SRC)

In [1], Wright et al. originally applied sparse presentation to FR and proposed the sparse representation based classifier (SRC). By

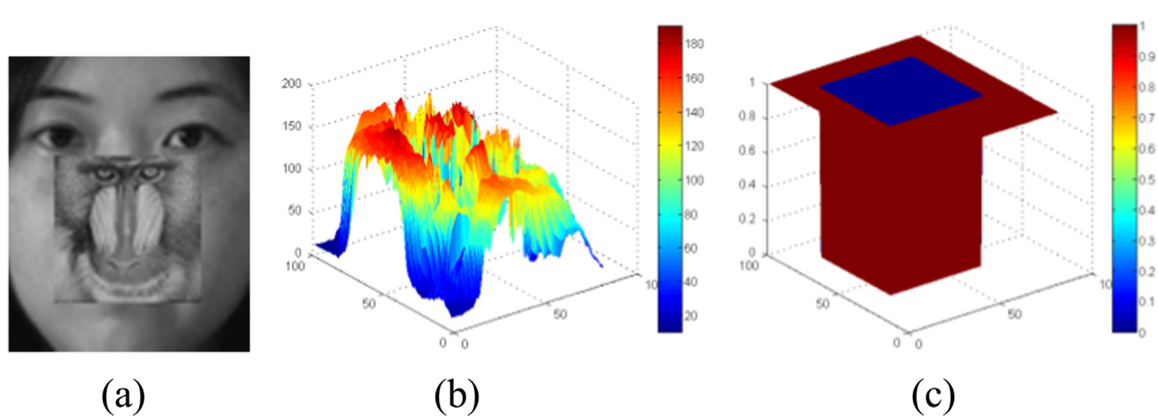

Fig. 1. The structured information of occluded image pixels. (a) a face image with occlusion; (b) pixels' values; (c) pixels' occlusion patterns with 1s indicating unoccluded pixels and 0s indicating occluded pixels. It is easy to see the occluded pixels' patterns but not the occluded pixels' values are spatially consistent.

Please cite this article as: X. Wang, et al., Structured regularized robust coding for face recognition, Neurocomputing (2016), http://dx.doi.org/10.1016/j.neucom.2016.06.065