Optik 125 (2014) 4729–4735

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

A hierarchical structure with improved OMP sparse representation used with face recognition

Jian Zhang[a,b,*], Hongzhi Zhang[c], Zhengming Li[d]

[a] Bio-computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, PR China
[b] Key Laboratory of Network Oriented Intelligent Computation, Shenzhen, PR China
[c] School of Computer Science and Technology, Harbin Institute of Technology, Harbin, PR China
[d] Guangdong Industry Training Centre, Guangdong Polytechnic Normal University, Guangzhou, PR China

ARTICLE INFO

Article history:
Received 10 September 2013
Accepted 24 April 2014

Keywords:
Hierarchical structure
Orthogonal matching pursuit
Sparse representation
Image classification

ABSTRACT

With the rapid development of the face recognition technology, more and more optical products are applied in people's real life. The recognition accuracy can be improved by increasing the number of training samples, but the colossal training samples will result in the increase of computational complexity. In recent years, sparse representation method becomes a research hot spot on face recognition. In this paper we propose an energy constrain orthogonal matching pursuit (ECOMP) algorithm for sparse representation to select the few training samples and a hierarchical structure for face recognition. We filter the training samples with ECOMP algorithm and then we compute the weights by all selected training samples. At last we find the closest recovery sample to the test sample. Simultaneously the experimental results in AR, ORL and FERET database also show that our proposed method has better recognition performance than the LRC and SRC_OMP method.

© 2014 Elsevier GmbH. All rights reserved.

1. Introduction

Biometrics has become one of the most important branches of pattern recognition, and the face recognition is one of the most attractive biometrics technologies. Face recognition is not only commonly used in people's daily life to confirm the identity, but also it has become the most popular pattern recognition research topics. Although the accuracy of face recognition is lower than iris recognition and fingerprint recognition, it is one of the most friendly biometric identification technologies. People do not reject because of its contactless and non-invasive. In recent years, a mass of study has focused on face recognition [1–5]. However, face recognition is still a great challenge owing to various illuminations, facial expression, face position and a variety of other environmental factors effects (Fig. 1).

Many of the face recognition methods have been reported, such as holistic-based feature recognition [6–9], matching by structural-based features [10–12] as well as methods based on hybrid features [13–15]. The first approach is based on facial geometry characteristics (geometric feature based), but it has low efficiency on identification. The principal component analysis (PCA) method is proposed later for face recognition [16]. PCA is mainly used for data dimensionality reduction. It discards the little changes dimension and lefts "fine" features, at the same time it makes smaller computation. Belhumeur et al. [17] adopt the linear discriminant analysis (LDA) method on the basis of the PCA. It transforms the principal component reduced the dimensionality in order to obtain the largest between-class scatter and inter-class scatter. This method is still the mainstream of face recognition method and produces a lot of different variants [18–29]. Another important method of face recognition is elastic graph matching (EGM) techniques [30], which uses an attribute graph to describe human face. The method does not only preserve the global structure of the facial features, but also it models key parts of the human face features. In recent years, there has been some further expansion of the method [31–33].

Lately, in the wake of widely application on compressed sensing (CS) technology, sparse representation research on face recognition technology is gradually increasing [34–40]. It is the important prerequisite of CS theory that how to find sparse representation for reconstructed sample. It directly affects the signal reconstruction accuracy. Literature [41] extended the CS theory from the orthogonal linear space to the redundant dictionary. It could obtain the samples of redundant dictionary by some algorithms from

* Corresponding author at: Bio-computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, PR China.
E-mail addresses: zpower007@163.com (J. Zhang), zhanghz@hit.edu.cn (H. Zhang), Lizhengming2004@126.com (Z. Li).

http://dx.doi.org/10.1016/j.ijleo.2014.05.018
0030-4026/© 2014 Elsevier GmbH. All rights reserved.

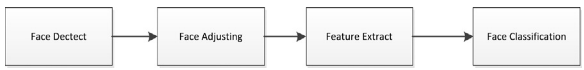

Fig. 1. The flow of the face recognition.

a few random observations. In the field of pattern recognition, the literature [42] proposed a classification method based on sparse reconstruction with CS theory (SRC). Its efficiency is better than the current Face recognition algorithm in the noiseless and noisy situations. Literature [43] studied the sparse representation theory and improved SRC method by quickly implement of matrix Cholesky decomposition algorithm. The method also achieved good recognition results, but the algorithm is sensitive to expression changes. Literature [44] replaced L1-norm minimization reconstruction algorithm with orthogonal matching pursuit (OMP) algorithm to reduce the computational complexity. In addition, it imported class information of each model to OMP. The proposed Classified Orthogonal Matching Pursuit (COMP) algorithm improved the classification performance.

In this paper, we use the hierarchical processing structure. There are three layers in this structure. In the first layer we select training samples with ECOMP algorithm. The algorithm exploits the orthogonally of projection OMP to select the training samples. The iteration process will stop when the residual is smaller than the energy threshold value. So that we can reconstitute a new sample training sequence. And then, in the second layer, we represent the test sample with reconstituted training samples in term of linear combination. At last, in the third layer, we recover the test sample with the linear combination of selected samples which are homogeneous class. The selected samples are obtained from the first layer and the weights are computed from the second layer. Each class existed in selected samples has a recovered test sample. We calculate the closest recovered sample to the test sample. The corresponding class is the final result by our proposed method.

Later in the paper, the structure is as follows: Section 2, we describe the proposed method. Section 3, we analyze the proposed method. In Section 4, we simulate the method based on database and analyze the experimental results. The last section we conclude the paper.

2. The description of ECOMP algorithm

We give an over complete dictionary matrix $A \in R^{n \times k}$, in which each column is an atoms of prototype. Given a test sample y, it can be expressed as a sparse linear combination of these atoms. Test sample y can be expressed as $y = Ax$ or $y \approx Ax$, satisfying $\left\| y - Ax \right\|_p \le \varepsilon$. "Over completing" means the number of atoms is far greater than the dimensionality of the testing sample y. Namely $n \ll k$.

The idea of matching pursuit is that we select an atom which is best matching to test sample y among the dictionary matrix, then we build a sparse approximation to calculate the residual signal. In the next step we continue to select the atom which is the best match to the residual signal iteratively. Test sample y can be equal to linear sum of these atoms together with the final residual. After K-step decomposition, we can get

$$y = \sum_{n=0}^{k} \left\langle R_n y, x_{r_n} \right\rangle R_n y + R_{k+1} y \tag{1}$$

where $R_0 y = y$. Or

$$y = \sum_{i=1}^{k} a_i x_{n_i} + R_k y = y_k + R_k y \tag{2}$$

where y_k is the optimal k-item approximation, if and only if $R_k y \in V_k^{\perp}$.

However, the vertical projection of a sample (residual) on the selected atom is non-orthogonality. This will bring about the results that each iteration is not optimal but sub-optimal. In addition, there are a lot of iterations for convergence. Accordingly, we need to orthogonally process the atom selected in each step in order to get faster convergence in the same accuracy.

We define

$$y = \sum_{n=1}^{k} a_n^k x_n + R_k y \quad n = 1, \ldots, k \tag{3}$$

(3) satisfies $\left\langle R_k y, x_n \right\rangle = 0$. $k+1$ term model is as follows:

$$y = \sum_{n=1}^{k+1} a_n^{k+1} x_n + R_{k+1} y \quad n = 1, \ldots, k+1 \tag{4}$$

(4) satisfies $\left\langle R_k f, x_n \right\rangle = 0$. We make $k+1$ term model to minus k term model, the result is the following:

$$\sum_{n=1}^{k} (a_n^{k+1} - a_n^k) x_n + a_{k+1}^{k+1} x_{k+1} + R_{k+1} y - R_k y \tag{5}$$

Because the atoms of dictionary matrix A are not orthogonal, we are necessary to introduce an auxiliary model depending on all of k terms $x_n (n = 1, \ldots, k)$, the x_{k+1} is described below:

$$x_{k+1} = \sum_{n=1}^{k} b_n^k x_n + \gamma_k \quad n = 1, \ldots, k \tag{6}$$

(6) satisfies $\left\langle \gamma_k, x_n \right\rangle = 0$. Where x_{k+1} is the one of the orthogonal projection operator on $span(x_1, \ldots, x_k)$, the latter term is residual. I.e. $\sum_{n=1}^{k} b_n^k x_n = P_{V_k} x_{k+1}$ and $\gamma_k = P_{V^{\perp}} x_{k+1}$.

We make (5) into (4), the result will be following:

$$\sum_{n=1}^{k} (a_n^{k+1} - a_n^k + a_{k+1}^{k+1} b_n^k) x_n + (a_{k+1}^{k+1} \gamma_k + R_{k+1} y - R_k y) = 0 \tag{7}$$

If

$$a_n^{k+1} - a_n^k + a_{k+1}^{k+1} b_n^k = 0 \tag{8}$$

And

$$a_{k+1}^{k+1} \gamma_k + R_{k+1} y - R_k y = 0 \tag{9}$$

We define

$$a_{k+1}^{k+1} = \alpha_k$$

So we can rewrite (8) into

$$a_n^{k+1} - a_n^k + \alpha_k b_n^k = 0 \quad n = 1, \ldots, k \tag{10}$$

where

$$\alpha_k = \frac{\left\langle R_k y, x_{k+1} \right\rangle}{\left\langle \gamma_k, x_{k+1} \right\rangle} = \frac{\left\langle R_k y, x_{k+1} \right\rangle}{\left\| \gamma_k \right\|^2}.$$

We make inner product on both sides in (9) with x_{k+1}. We will obtain the first portion of a^k by solve. Similarly we make inner product on both sides in (5) with γ_k to obtain the second portion of a^k by solve.

However, a set of training samples (atoms) number is often limited for face recognition. It is difficult to build a complete dictionary database. Therefore, we need to design a convergence condition to terminate the iterative, at the same time we will also ensure decomposition of the residual $R_l y$ as small as possible.

In this paper, we design the residual energy threshold method as iteration termination condition. And through image restoration ideas, we try to ensure that the result by limited training sample is as soon as possible to the test sample. So we will get new sub-samples extracted from all training samples.

According to ECOMP algorithm, the input test sample y can be obtained for a sparse representation:

$$y = \sum_{n=1}^{l} a_n^l x_n + R_l y \quad n = 1, \ldots, l \tag{11}$$

Download English Version:

https://daneshyari.com/en/article/849441

Download Persian Version:

https://daneshyari.com/article/849441

Daneshyari.com