Engineering Applications of Artificial Intelligence 34 (2014) 13–22

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

A novel local extrema based gravitational search algorithm and its application in face recognition using one training image per class

Tapabrata Chakraborti [a], Kaushik Das Sharma [b], Amitava Chatterjee [a,*]

[a] Department of Electrical Engineering, Jadavpur University, Kolkata 700032, India
[b] Department of Applied Physics, University of Calcutta, Kolkata 700009, India

ARTICLE INFO

Article history:
Received 23 March 2013
Received in revised form
1 December 2013
Accepted 5 May 2014
Available online 2 June 2014

Keywords:
Gravitational search algorithm (GSA)
Fisher linear discriminant analysis (FLDA)
Singular value decomposition (SVD)
2-D random local extrema gravitational search algorithm (2-D RLEGSA)

ABSTRACT

In this present paper a new methodology has been presented involving a stochastic optimization based approach to solve the face recognition problem with only one training image per class. Singular value decomposition (SVD) is used to decompose the single training image into two component images in order to compute the within class scatter matrix. The stochastic optimization approach is implemented employing gravitational search algorithm (GSA) which searches for an optimal transform matrix instead of using the traditional solution of general eigenvalue problem as is carried out in Fisher linear discriminant analysis (FLDA). The present paper also proposes two novel variants of GSA, namely the 2-D version of GSA, in order to cater for the 2-D image data, and the other one is a 2-D randomized local extrema based GSA (RLEGSA), which employs a stochastic local neighborhood based search instead of global search, as in basic GSA. Finally, a novel concept of performing an automated selection of projection vectors is incorporated in the 2-D RLEGSA to propose an improved variant, called the Modified RLEGSA (MRLEGSA). Experimental results, based on benchmark Yale A and ORL databases, show that the proposed methods outperform several existing schemes.

1. Introduction

In recent years face recognition has become a widely researched topic since it has numerous real world applications like authentication, identification, advanced human computer interaction and many other emerging fields of research. Face recognition spans the subjects of pattern recognition, image processing, computer vision, machine learning, etc. With the growing importance of biometric recognition systems (Jain and Prabhakar, 2004), due to low susceptibility to security loss, face recognition based biometrics has gained much popularity in recent times.

Many approaches to face recognition have been proposed over the last two decades (Zhao et al., 2003; Jafri and Arabnia, 2009; Chakrabarty et al., 2013) most of which are based on supervised learning. Hence they follow a common sequence of steps. There is a feature extraction step in which a set of discriminating features are extracted from a set of training images (Brunelli and Poggio, 1993). Then if the set of features extracted is dimensionally large, there may be a feature selection/reduction procedure where a reduced set of highly discriminating features are selected employing a suitable algorithm which may attempt to optimize a suitable cost function. For the face recognition problem, since images of different persons are, after all, human faces they have some common characteristics which indicates that some features in a large set of extracted features will not have enough discriminating power. This makes feature selection/reduction an important step as a large feature set might not necessarily result in a higher recognition rate (Tu et al., 2007). This step is followed by the classification step where the final conclusion regarding recognition or authentication is actually performed. Several variations of methods proposed in each of these steps generates new approaches in solving the problem.

A general drawback of the supervised learning method is that for a good classification accuracy rate, the number of training samples needs to be sufficiently large (depending on the number of test images and the number of classes). In those particular methods where inter class and intra class distances are used, the methods do not work at all when there is a single training image of each subject because in this case, the intra class distances are not defined (Gao et al., 2008). This drawback is prominent in several approaches which include the popular methodology of Fisher linear discriminant analysis (FLDA). A few methods have been proposed in recent years to solve this problem of FLDA based face recognition where there is only one training image per person e.g. generalized inverse method (Tian et al., 1988), perturbation based method (Hong and Yang 1991), direct FLDA method (Yu and Yang 2001), null space method (Lu et al., 2003), 2-D FLDA method

* Corresponding author. Tel.: +91 3324146949.
E-mail address: cha_ami@yahoo.co.in (A. Chatterjee).

http://dx.doi.org/10.1016/j.engappai.2014.05.002
0952-1976/

(Ye et al., 2004), singular value decomposition (SVD) based method (Gao et al., 2008) etc.

Our present research concentrates on those more challenging face recognition problems which suffer from small sample size (SSS) problem, typically in those situations where there is only a single training sample available per class/person. The research on face recognition problem using single training sample per person is well known as a very challenging problem and it has gained prominence in recent times (Gao et al., 2008; Tan et al., 2006; Zhu et al., 2012). This situation arises in many real-world scenarios such as utilization of smart cards, airport check-in and check-out, special situations of law enforcement, critical surveillance scenarios and checking for access control etc. (Tan et al., 2006; Zhu et al., 2012). Our present work is inspired by the method proposed by Gao et al. (2008) in which the single training image of a particular class is decomposed into two component images using SVD and then the intra class distance can be conveniently determined using these two resulting images. However, the SVD based FLDA approach in Gao et al. (2008) uses the general eigen value theory to solve the cost function. In this paper we propose a novel method of solving the FLDA cost function using an intelligent iterative stochastic optimization algorithm which can simultaneously solve the feature selection/reduction phase along with the feature extraction phase, thus effectively merging the operations required in two steps. The iterative stochastic optimization problem is solved using a recently proposed method, called gravitational search algorithm (GSA). GSA is a powerful iterative optimization algorithm based on Newton's laws of gravity and motion (Rashedi et al., 2009; Pal et al., 2013). Several interesting applications have recently been proposed using GSA in the domains of e.g. image processing (Sun and Zhang, 2013) and data clustering (Hatamlou et al., 2012; Hatamlou et al., 2011). Three modifications of the GSA have been proposed in this paper with the objective of solving our problem. The first variation proposes a 2-D GSA to adopt the GSA in processing of 2-D images. The second variation introduces a novel random local extrema based GSA (RLEGSA). To the best of our knowledge and belief, although some local best methods have been proposed earlier for a similar swarm intelligence based method called particle swarm optimization (PSO) (Suganthan, 1999; Das Sharma et al., 2012), this is the first such variation developed in the genre of GSA. The third variation incorporates the automated selection of projection vectors within the GSA based cost function optimization framework.

The rest of this paper is organized as follows. Section 2 provides a description of the SVD and FLDA based feature extraction schemes for the single training image per person scenario. Section 3 describes an overview of the traditional gravitational search algorithm and detailed descriptions of the novel variants proposed in this work. Section 4 presents the experiments and simulation results. Section 5 concludes the paper.

2. SVD and FLDA based feature extraction scheme

Let us consider there are C classes with each having a single image $I_k \in \mathfrak{R}^{m\times n}$ (k=1,…,C). If $m \geq n$, then let $U_k \in \mathfrak{R}^{m\times m}$ and $V_k \in \mathfrak{R}^{n\times n}$ be the eigenvector matrices of $I_k I_k^T$ and $I_k^T I_k$ respectively. Let u_i^k and v_i^k be the ith column of U_k and V_k respectively. Let σ_i^k be the ith singular value of I_k such that σ_i^k is in descending order of magnitude as i increases i.e. $\sigma_1^k \geq \sigma_2^k \geq \cdots \sigma_{i-1}^k \geq \sigma_i^k \geq \sigma_{i+1}^k \cdots \sigma_n^k$. Then the image can be described as (Gao et al., 2008; Golub and Loan, 1983)

$$I_k = \sum_{i=1}^{n} \sigma_i^k u_i^k (v_i^k)^T \tag{1}$$

Hence each image can be thought of being constituted as a summation of n basis images where each basis image corresponds to a particular singular value and the energy content of a basis image is higher if its associated singular value is higher in magnitude. Following the philosophy described in Gao et al. (2008), an image $\hat{I}_k$ is constructed taking the three most significant SVD basis images as

$$\hat{I}_k = \sum_{i=1}^{3} \sigma_i^k u_i^k (v_i^k)^T \tag{2}$$

Thus after obtaining $\hat{I}_k$ we have two image matrices I_k and $\hat{I}_k$ in each class k, and we can also compute a difference image, $\Delta I_k = I_k - \hat{I}_k = \sum_{i=4}^{n} \sigma_i^k u_i^k (v_i^k)^T$. The creation of $\hat{I}_k$ and ΔI_k facilitates the calculation of an approximate within-class scatter matrix, which is, otherwise, not possible when we have only one training image per person/class. Let the within class scatter matrix be denoted as S_w and the between class scatter matrix be denoted as S_b. To compute S_w and S_b we need to compute the global mean image $\bar{I}$ and mean image of the kth class $\bar{I}_k$, which are defined as (Gao et al., 2008)

$$\bar{I}_k = \tfrac{1}{2}(I_k + \hat{I}_k) \tag{3}$$

$$\bar{I} = \frac{1}{2C} \sum_{k=1}^{C} (I_k + \bar{I}_k) \tag{4}$$

Then S_w and S_b can be computed as

$$S_b = \frac{1}{C} \sum_{k=1}^{C} (\bar{I}_k - \bar{I})^T (\bar{I}_k - \bar{I}) \tag{5}$$

$$S_w = \frac{1}{C} \sum_{k=1}^{C} [(I_k - \bar{I}_k)^T (I_k - \bar{I}_k) + (\hat{I}_k - \hat{I}_k)^T (\hat{I}_k - \bar{I}_k)] \tag{6}$$

Using (3) and (4), one can obtain (Gao et al., 2008)

$$S_w = \frac{1}{2C} \sum_{k=1}^{C} (I_k - \hat{I}_k)^T (I_k - \hat{I}_k) \tag{7}$$

From the theory of two-dimensional FLDA (Ye et al., 2004), then our goal will be to seek a set of d optimal discriminating column vectors w_j (j=1, 2, …, d) constituting $m \times d$ optimal projection matrix W so as to minimize the cost function

$$J(W) = \frac{trace(W^T S_w W)}{trace(W^T S_b W)} \tag{8}$$

Once W is determined, all the training images are projected on to W to obtain the feature matrices Z_k

$$Z_k = I_k \times W, \quad k = 1, 2, \ldots, C \tag{9}$$

Then, for each class k we have a feature matrix Z_k. If we have an input test image I then the corresponding feature matrix is Z=$I \times W$

Then we can utilize the nearest neighbor classifier method, where we calculate the Euclidean distance D_k of feature matrix Z from each feature matrix Z_k pertaining to class k, given as

$$D_k = \|Z_k - Z\|, \quad k = 1, \ldots, c \tag{10}$$

Then, I is identified to belong to that class k for which D_k is minimum.

Now in Gao et al. (2008) the optimal value of W has been determined by using the general eigen value theorem and for that the cost function considered is the inverse of (8), which is the traditional cost function considered in Ye et al. (2004), as they solved a maximization problem. In this work, as we attempt to solve a minimization function by utilizing the gravitational search algorithm and its proposed variants, we have utilized a form of $J(W)$ which is inverse of that considered in Ye et al. (2004) and Gao et al. (2008).