Neurocomputing 86 (2012) 184–192

Contents lists available at SciVerse ScienceDirect

Neurocomputing

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

Letters

An affine invariant discriminate analysis with canonical correlation analysis

Rushi Lan [a], Jianwei Yang [a,*], Yong Jiang [a], Zhan Song [b,c], Yuan Yan Tang [d]

[a] College of Math and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China
[b] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China
[c] The Chinese University of Hong Kong, Hong Kong, China
[d] Department of Computer Science, Chongqing University, Chongqing 400030, China

ARTICLE INFO

Article history:
Received 16 September 2011
Received in revised form
9 January 2012
Accepted 19 January 2012
Communicated by C. Fyfe
Available online 1 March 2012

Keywords:
Affine invariant discriminate analysis (AIDA)
Canonical correlation analysis (CCA)
Central projection transform (CPT)
Affine transformation

ABSTRACT

Canonical correlation analysis (CCA) is invariant with regard to affine transformation, but it cannot be directly applied to affine invariant pattern recognition. The reason mainly lies in that many existing CCA-based schemes represent the pattern by matrix-to-vector method, as a result, the structure and spatial information of the original pattern is discarded. In this paper, an affine invariant discriminate analysis (AIDA) method is developed for pattern recognition. Dislike the matrix-to-vector representation, an object is first converted to a projection matrix by central projection transform (CPT). After a point matching process, CCA is performed to projection matrices of the object and the model, and two vectors will be derived. Therefore, the object is classified to a model by the smallest distance between the obtained vectors. Comparisons of experimental results are given with respect to some existing methods, which demonstrate the effectiveness of the proposed AIDA method.

1. Introduction

Pattern recognition [1] is the study of how machines can observe the environment, learn to distinguish patterns of interest from their background, and make sound and reasonable decisions about the categories of the patterns. A large number of methods have been proposed to deal with various pattern recognition tasks. Of the existing schemes, the statistical approach has been most intensively investigated and applied in many real-life situations. Principal component analysis (PCA) [2] and linear discriminant analysis (LDA) [3] are two most representative methods. Lots of new approaches, such as 2DPCA [4], exponential discriminant analysis (EDA) [5] etc., have been proposed to improve the performance of traditional methods.

Recently, a multivariate statistical analysis technique, canonical correlation analysis (CCA), arouses a great deal of interest among researchers. CCA, proposed by Hotelling [6], measures linear relationships between two multidimensional variables. It finds base vectors (canonical factors) for two variables such that the correlations between the projections of the variables onto these canonical factors are mutually maximized. The directions of canonical factors capture functional relations of the two variables. Comparing with other two commonly used multivariate statistical analysis techniques, namely PCA and LDA, CCA not only reduces the dimensionality of the original variables but also seeks directions for two variables to maximize their correlations. Therefore, it may be better suited for some recognition tasks. Many promising results have been achieved based on CCA. Sun et al. [7] apply CCA to feature fusion and image recognition for the first time. They propose a feature fusion method which uses correlation feature of two groups of feature as effective discriminant information. Sun et al. propose locality preserving CCA [8] and sample label-based CCA [9] methods, which greatly improve the limitation of the original CCA.

Another property of CCA [10] is that it is invariant with regard to affine transformation. Affine invariant pattern recognition, extracting invariant features from the objects under affine transformation, plays an important role in object recognition, and has found successful applications in many fields. Many algorithms have been developed for affine invariant pattern recognition. Of the existing methods, wavelet descriptors [11], affine moment invariants (AMIs) [12,13], and Multi-Scale Autoconvolution (MSA) [14], etc. are some representative approaches. However, very few attention has been paid on the affine invariant property of CCA. The reason may lie in that almost every existing CCA-based method, without exception, reshapes the original two dimensional image into a vector. This matrix-to-vector operation leads to two following problems in affine invariant pattern recognition:

- Firstly, the size of the image may be changed as a result of the affine transformation. CCA cannot deal with vectors obtained from images with different sizes. This problem can be overcome

* Corresponding author.
E-mail address: yjianw2002@yahoo.com.cn (J. Yang).

0925-2312/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2012.01.026

by normalizing the image into the same size, but this way may effect the accuracy of the image.

- Secondly, the structure and spatial information of the image matrix cannot be maintained in the vector. If the image is modeled as a high-dimensional vector, it is very difficult to yield a reliable and robust estimation of the underlying data distribution. For example, we consider the following two matrices and their vector representations:

$$\begin{bmatrix} 0 & 1 & 0 \\ 4 & 5 & 2 \\ 0 & 3 & 0 \end{bmatrix} \mapsto [0\ 1\ 0\ 4\ 5\ 2\ 0\ 3\ 0]$$

and

$$\begin{bmatrix} 0 & 2 & 0 \\ 1 & 5 & 3 \\ 0 & 4 & 0 \end{bmatrix} \mapsto [0\ 2\ 0\ 1\ 5\ 3\ 0\ 4\ 0].$$

We can find that the two matrices are connected with a rotation transformation by observing the data distribution in the matrix. On the other hand, we cannot intuitively find the relation between the two vectors. The relation of the two vectors is not obvious as that represented in matrix form.

Therefore, CCA cannot directly be applied to affine invariant shape recognition.

In this paper, an affine invariant discriminate analysis (AIDA) method based on CCA and central projection transform (CPT) [15] is proposed. Consisting of two stages, namely CPT and CCA, AIDA can be formally regarded as following form: AIDA=CPT+CCA. In order to address the aforementioned problems, CPT and a point matching process are carried out before CCA respectively. Using CPT, any object can be projected to a closed curve, which preserves the affine transformation and is called general contour (GC) of the object, and the image is converted to a $2 \times N$ matrix whose column vectors are the coordinate value of GC. As a result, images with different sizes are transformed to the matrices with the same sizes. And then a point matching process is carried out to the obtained matrix to preserve the structure and spatial information of the image. After that, CCA is employed to classify objects under affine transformation. Several experiments have been conducted to evaluate the performance of AIDA, and satisfying results are obtained.

The rest of the paper is organized as follows: some preliminaries are introduced in Section 2. In Section 3, the proposed AIDA method is given. The experiments and results are shown in Section 4. In the last section, some conclusions will be given.

2. Preliminaries

2.1. Affine transformation

Considering a parametric point $X(t)=[x(t),y(t)]'$ with the parameter t of a region, then the affine transformation consists of a linear transformation and translation as follows:

$$\tilde{x}(t)=a_{11}x(t)+a_{12}y(t)+b_1,$$

$$\tilde{y}(t)=a_{21}x(t)+a_{22}y(t)+b_2.$$

The above equations can be written with the following form:

$$\begin{bmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = A \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + B, \quad (1)$$

where the non-singular matrix A represents the scaling, rotation, skewing transformations and the vector B corresponds to the translation.

2.2. CPT

In this subsection, the CPT method and its properties are introduced. We have proposed CPT in [15], which can be considered as a sub-case of the Radon transform [16,17], and further developed in [18,19] to extract rotation invariant features. The affine invariant property of CPT is studied in this paper. In CPT, a closed contour can be derived from the object by taking projection along lines from the centroid with different angles.

Suppose that an object in the 2D plane is represented by $I(x,y)$. To perform CPT, the Cartesian coordinate system should firstly be converted to polar coordinate system. For achieving the translation invariance, the pole of polar coordinate system is taken at the centroid of the object which can be computed from the first geometric moments. Let $f(r,\theta)$ be the transformed image. After the conversion of the system, the CPT is performed by computing the following integral:

$$g(\theta) = \int f(r,\theta)\, dr, \quad (2)$$

where $\theta \in [0, 2\pi]$. $(g(\theta),\theta)$ denotes a point in the plane of R^2. Let θ goes from 0 to 2π, then the $\{(g(\theta),\theta) | \theta \in [0, 2\pi]\}$ is a closed curve, which is the GC of the object.

From a practical point of view, the images to be analyzed by a recognition system are most often stored in discrete formats. Catering to such two-dimensional discrimination patterns, we should modify Eq. (2) into the following expressions:

$$g(\theta_n) = \sum_{m=0}^{M-1} f(r_m,\theta_n)\, dr, \quad (3)$$

where $\theta_n \in [0, 2\pi]$, $n=0,1,2,\ldots,N-1$. M and N in Eq. (3) denote the sampling intervals applied in the coordinate system conversion produce.

Based on the GC, the original object can be converted to the following $2 \times N$ matrix using coordinate conversion

$$\begin{bmatrix} g(\theta_0)\cos\theta_0 & g(\theta_1)\cos\theta_1 & \cdots & g(\theta_{N-1})\cos\theta_{N-1} \\ g(\theta_0)\sin\theta_0 & g(\theta_1)\sin\theta_1 & \cdots & g(\theta_{N-1})\sin\theta_{N-1} \end{bmatrix}. \quad (4)$$

This matrix is called the projection matrix of the object, whose column vectors are the Cartesian coordinate values of the GC.

It can be proved that GC derived from the affine transformed object is the same affine transformed version of GC derived from the original object. For example, two images shown in Fig. 1(a) and (b) are related by an affine transformation. The GCs of the them are plotted in Fig. 1(c) and (d) respectively. From Fig. 1, we can observe that the CPT method is able to keep the affine transformed information. Besides, images with different sizes (see Fig. 1(a) and (b)) are converted to GCs with the same length by CPT.

2.3. CCA

Considering two zero-mean random variables $\xi \in R^p$ and $\zeta \in R^q$, CCA aims to find a pair of basis or projection vectors, ω_ξ and ω_ζ, which maximize the correlation between the projections $\alpha = \omega_\xi^T \xi$ and $\beta = \omega_\zeta^T \zeta$. More formally, the basis vectors can be formulated as following:

$$(\omega_\xi,\omega_\zeta) = \arg\max_{\omega_\xi,\omega_\zeta} \frac{E[\alpha\beta]}{\sqrt{E[\alpha^2]E[\beta^2]}} = \arg\max_{\omega_\xi,\omega_\zeta} \frac{E[\omega_\xi^T \xi\zeta^T \omega_\zeta]}{\sqrt{E[\omega_\xi^T \xi\xi^T \omega_\xi]E[\omega_\zeta^T \zeta\zeta^T \omega_\zeta]}}$$

$$= \arg\max_{\omega_\xi,\omega_\zeta} \frac{\omega_\xi^T C_{\xi\zeta}\omega_\zeta}{\sqrt{\omega_\xi^T C_{\xi\xi}\omega_\xi \omega_\zeta^T C_{\zeta\zeta}\omega_\zeta}}, \quad (5)$$

where $C_{\xi\xi} \in R^{p\times p}$ and $C_{\zeta\zeta} \in R^{q\times q}$ are the non-singular within-set covariance matrices of x and y, while $C_{\xi\zeta} \in R^{p\times q}$ denotes their between-set covariance matrix.