IDEA GROUP PUBLISHING
701 E. Chocolate Avenue, Suite 200, Hershey PA 17033-1240, USA
Tel: 717/533-8845; Fax 717/533-8661; URL-http://www.idea-group.com

ITB12297

This paper appears in the publication, *Biometric Image Discrimination Technologies*
by David Zhang, Xiaoyuan Jing, and Jian Yang © 2006, Idea Group Inc.

Chapter II

Principal Component Analysis

ABSTRACT

In this chapter, we first describe some basic concepts of PCA, a useful statistical technique that can be used in many fields, such as face patterns and other biometrics. Then, we introduce PCA definitions and related technologies. Following, we discuss non-linear PCA technologies. Finally, some useful conclusions are summarized.

INTRODUCTION

PCA is a classical feature extraction and data representation technique widely used in pattern recognition and computer vision (Duda, Hart, & Stork, 2000; Yang, Zhang, Frangi, & Yang, 2004; Anderson, 1963; Kim, n.d.; Boser, Guyon, & Vapnik, 1992). Sirovich and Kirby first used PCA to efficiently represent pictures of human faces (Sirovich & Kirby, 1987; Kirby & Sirovich, 1990). They argued that any face image could be reconstructed approximately as a weighted sum of a small collection of images that define a facial basis (eigenimages) and a mean image of the face. Since eigenpictures are fairly good at representing face images, one could consider using the projections along them as classification features for recognizing human faces. Within this context, Turk and Pentland presented the well-known eigenfaces method for face recognition in 1991 (Turk & Pentland, 1991). They developed the well-known face recognition method, where the eigenfaces correspond to the eigenvectors associated with the dominant eigenvalues of the face covariance matrix. The eigenfaces define a feature space, or “face space,” which drastically reduces the dimensionality of the original space, and face detection and identification are carried out in the reduced space (Zhang, 1997). Since then, PCA has been widely investigated and become one of the most successful approaches in face recognition (Pentland, 2000; Grudin, 2000; Cottrell & Fleming, 1990; Valentin, Abdi, O’Toole, & Cottrell, 1994). Penev and Sirovich discussed the problem of the dimensionality of the “face space” when eigenfaces are used for representation (Penev & Sirovich, 2000). Zhao and Yang tried to account for the arbitrary effects of illumination in PCA-based vision systems by generating an analytically closed form formula of the covariance matrix for the case with a special lighting condition and then generalizing to an arbitrary illumination via an illumination equation (Zhao & Yang, 1999). However, Wiskott, Fellous, Krüger and von der Malsburg (1997) pointed out that PCA could not capture even the simplest invariance unless this information is explicitly provided in the training data. They proposed a technique known as elastic bunch graph matching to overcome the weaknesses of PCA. In this chapter, we will show you some basic definitions of PCA.

DEFINITIONS AND TECHNOLOGIES

Mathematical Background of PCA

This section will attempt to give some elementary background mathematical skills required to understand the process of PCA (Smith, 2002; Vapnik, 1995).

Eigenvectors and Eigenvalues

Given a d-by-d matrix $\mathbf{M}$, a very important class of equation is of the form (Duda, Hart, & Stork, 2000):

$$\mathbf{M}\mathbf{x} = \lambda\mathbf{x} \tag{2.1}$$

for scalar λ, which can be written:

$$(\mathbf{M} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \tag{2.2}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{0}$ is the zero vector. The solution vector $\mathbf{x} = \mathbf{e}_i$ and corresponding scalar $\lambda = \lambda$ are called the *eigenvector* and associated *eigenvalue*, respectively. If $\mathbf{M}$ is real and symmetric, there are d (possibly nondistinct) solution vectors $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_d\}$, each with an associated eigenvalue $\{\lambda_1, \lambda_2, \ldots, \lambda_d\}$. Under multiplication by $\mathbf{M}$ the eigenvectors are changed only in magnitude, not direction:

$$\mathbf{M}\mathbf{e}_j = \lambda_j\mathbf{e}_j \tag{2.3}$$

If $\mathbf{M}$ is diagonal, then the eigenvectors are parallel to the coordinate axes.

One method of finding the eigenvalues is to solve the *characteristic equation* (or *secular equation*):

Related Content