Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition 38 (2005) 997–1008
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

Visual object recognition using probabilistic kernel subspace similarity

Jianguo Lee[a], Jingdong Wang[b], Changshui Zhang[a,*], Zhaoqi Bian[a]

[a]*Department of Automation, Tsinghua University, Beijing 100084, P.R. China*
[b]*Department of Computer Science, Hongkong University of Science and Technology, Hongkong, P.R. China*

Received 6 April 2004; received in revised form 10 January 2005; accepted 10 January 2005

Abstract

Probabilistic subspace similarity-based face matching is an efficient face recognition algorithm proposed by Moghaddam et al. It makes one basic assumption: the intra-class face image set spans a linear space. However, there are yet no rational geometric interpretations of the similarity under that assumption. This paper investigates two subjects. First, we present one interpretation of the intra-class linear subspace assumption from the perspective of manifold analysis, and thus discover the geometric nature of the similarity. Second, we also note that the linear subspace assumption does not hold in some cases, and generalize it to nonlinear cases by introducing kernel tricks. The proposed model is named probabilistic kernel subspace similarity (PKSS). Experiments on synthetic data and real visual object recognition tasks show that PKSS can achieve promising performance, and outperform many other current popular object recognition algorithms.

Keywords: Subspace analysis; Tangent distance; Kernel methods; Visual object recognition; Manifold analysis

1. Introduction

Visual object recognition addresses the problem of finding and identifying objects in images. It is of fundamental importance to machine vision. For example, it is critical for a robot or an intelligent system to understand the environment [1]. In recent years, visual object recognition research has witnessed a growing interest in subspace analysis methods [2].

One classical subspace analysis approach is principal component analysis (PCA). In visual object recognition, PCA projects data onto a low-dimensional linear subspace in a minimum sum-squared error sense, then adopts these principal projected components as features for further recognition. However, PCA has at least one drawback: it regards the minor components as noise, and discards them in the following recognition procedure, which makes it an important issue to choose a proper dimension for the principal subspace. An intuitive solution is to choose the dimensionality according to the energy of PCA eigenvalues [3], such as choosing a dimensionality when the cumulative eigenvalue energy is greater than 90%. However, practice shows that applying this idea does not lead to performance improvement in many cases, while increases the computational cost remarkably sometimes.

Many researchers have noticed this problem in the field of object recognition, pointing out that information should be utilized not only from the principal subspace, but also from its orthogonal complemental subspace. In other words, it is favored to construct a noise-involved subspace model. There are two typical solutions towards this issue. One is

* Corresponding author. Tel.: +86 10 62782447; fax: +86 10 62786911.
E-mail address: zcs@mail.tsinghua.edu.cn (C. Zhang).

0031-3203/$30.00
doi:10.1016/j.patcog.2005.01.007

probabilistic PCA (PPCA) [4], which constructs a latent variable model over PCA to characterize noise, and adopts a maximum-likelihood (ML) method to estimate the noise part, which is also called minor component. The other work is by Moghaddam and Pentland [5], namely probabilistic subspace analysis (PSA), which has been successfully employed in visual learning field. PSA estimates two marginal Gaussian densities in those two subspaces, and forms a complete density representation by the product of these two independent marginal densities. In this way, the PSA model becomes a complete representation of the whole space. In fact, PSA is equivalent to PPCA in characterizing observed variables (a proof is given in the appendix).

In visual object recognition, the application of PSA leads to the probabilistic subspace similarity (PSS) algorithm [6,7], which is known as one of the best face recognition algorithms in the FERET face recognition test [8]. The important strategy adopted by PSS in face recognition is the intra-class face image subspace, which assumes that the difference between any two face images from the same subject spans a linear space. And the PSA model is employed to characterize this space. This assumption can be intuitively explained as ignoring the face variations both within and outside of the subspace. To provide an in-depth interpretation, Moghaddam also investigated PSS versus other so-called principal manifold techniques in the application of face recognition, including PCA, independent component analysis (ICA) and kernel PCA [7]. However, this work did not clearly present what kind of manifold PSS exactly represents. There are also some other theoretic interpretations from different perspectives. For example, Teh and Hinton [9] interpreted it from the perspective of Boltzmann machines, and made a nonlinear generalization in the framework of rate-coded restricted Boltzmann machines. Recently, Wang and Tang proposed a unified framework on PCA, linear discriminative analysis (LDA) and PSS in the field of face recognition [10]. They empirically considered that these three techniques characterize different kinds of face variations (intrinsic, transformation, noise). Although their interpretation, in some sense, touched the nature of the intra-class linear subspace assumption, it did not reveal the geometric meaning of the PSS measure yet. The first goal of this paper is to present a novel geometric interpretation to help us well understand the nature of the PSS algorithm.

In the procedure of interpretation, we also note that the linear subspace assumption does not hold in some data sets with complex structure. The other goal of this paper is straightforward: we extend the linear subspace assumption of PSS to nonlinear cases by introducing kernel tricks [11,12], so that the nonlinear extension could perform well in data sets with complex structure. The proposed algorithm is named probabilistic kernel subspace similarity (PKSS).

The rest of this paper is organized as follows. In Section 2, we briefly introduce the probabilistic subspace similarity and present a geometric interpretation. In Section 3, we propose the PKSS algorithm. In Section 4, we present experimental results on synthetic data set and real visual object recognition tasks. Finally, in Section 5, we draw conclusions.

2. Probabilistic subspace similarity and its geometric nature

In this section, we first briefly introduce PSA and PSS. Then, we present a geometric interpretation of the PSS similarity.

2.1. Probabilistic subspace analysis

Given a training image set $X = \{x_i\}_{i=1}^{m} \subset \mathscr{R}^d$, PCA can be adopted to characterize the image space. However, as having been pointed out previously, the principal eigensubspace by PCA is not a complete and good representation of the full image space.

The image space can also be represented by a Gaussian density function

$$P(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}(x-\mu)^{\mathrm{T}}\Sigma^{-1}(x-\mu)\right\}. \tag{1}$$

The dimension of the image space is usually comparable to or even larger than (i.e. $d > m$) the number of training image, which makes the second-order statistics (i.e. Σ) unreliable. Thus it is difficult to obtain the exact Gaussian density representation.

PSA is a probabilistic eigen-space representation algorithm towards both the shortcomings of PCA and the limitations of Gaussian density estimation. PSA divides the full image space X into principal subspace F and its orthogonal complemental subspace $\bar{F}$, which is illustrated in Fig. 1. Note that the principal subspace F can also be spanned by the first p principal components of PCA on X. Then, PSA estimates two Gaussian densities in these two subspaces, and the complete density estimation can be written as the

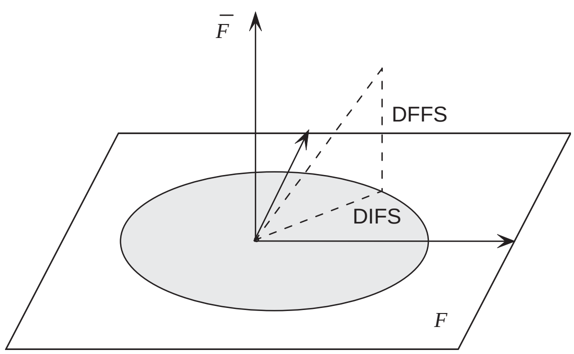

Fig. 1. Decomposition of the full image space X into principal subspace F and its orthogonal complemental subspace $\bar{F}$: $X = F \oplus \bar{F}$.