Neurocomputing 140 (2014) 104–113

Contents lists available at ScienceDirect

Neurocomputing

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

Feature extraction using local structure preserving discriminant analysis

Pu Huang [a,b], Caikou Chen [a,*], Zhenmin Tang [b], Zhangjing Yang [b]

[a] College of Information Engineering, Yangzhou University, Yangzhou 225009, China
[b] School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

A R T I C L E I N F O

Article history:
Received 16 July 2013
Received in revised form
3 March 2014
Accepted 10 March 2014
Communicated by Dr. J. Zhang
Available online 8 April 2014

Keywords:
Feature extraction
Manifold learning
LPP
CMVM
Local structure

A B S T R A C T

In this paper, an efficient feature extraction method, named local structure preserving discriminant analysis (LSPDA), is presented. LSPDA constructs the local scatter and the between-class scatter to characterize the sub- and multi-manifold information respectively. More specifically, the local structure is constructed according to a certain kind of similarity between data points which takes special consideration of both the local information and the class information based on a parameter-free neighborhood decision rule, and the between-class structure is constructed according to the importance degrees of the not-same-class points measured by a strictly monotonically decreasing function. After the local scatter and the between-class scatter have been characterized, the novel feature extraction criterion is derived via maximizing the difference between the between-class scatter and the local scatter. Experimental results on the Wine dataset, AR, FERET, CMU PIE, ORL and LFW face databases show the effectiveness of the proposed method.

1. Introduction

Feature extraction from the original data, which is often dictated by practical feasibility, is an important research topic in computer vision and pattern recognition. Over the past few decades, a number of useful feature extraction methods have been developed. Principal component analysis (PCA) [1] and linear discriminant analysis (LDA) [2] are the two most well-known methods for linear feature extraction.

PCA is an unsupervised learning algorithm, which projects the original data into a low-dimensional subspace spanned by the eigenvectors associated with the largest eigenvalues of the data's covariance matrix. Unlike PCA, LDA is a supervised method, which aims to find an optimal projection by maximizing the ratio of the trace of the between-class scatter to the trace of the within-class scatter. Due to encoding the class label information of the data, LDA has a more discriminative ability than PCA. Despite the success of LDA in many applications, its effectiveness is still limited because the number of the available projection directions is lower than the class number. Furthermore, LDA cannot be applied directly to small sample size (SSS) problems [3] because the within-class scatter matrix is singular. In the past, many LDA extensions [4–14] have been developed to deal with this problem, such as Pseudo-inverse LDA (PLDA) [4], regular LDA (RLDA) [5], penalized discriminant analysis (PDA) [6], LDA/GSVD [7], LDA/QR [8], orthogonal LDA (OLDA) [9], null space LDA (NLDA) [10], direct LDA (DLDA) [11], CLDA [12], two-stage LDA [13], maximum margin criterion (MMC) [14]. In [15], considering the performance of

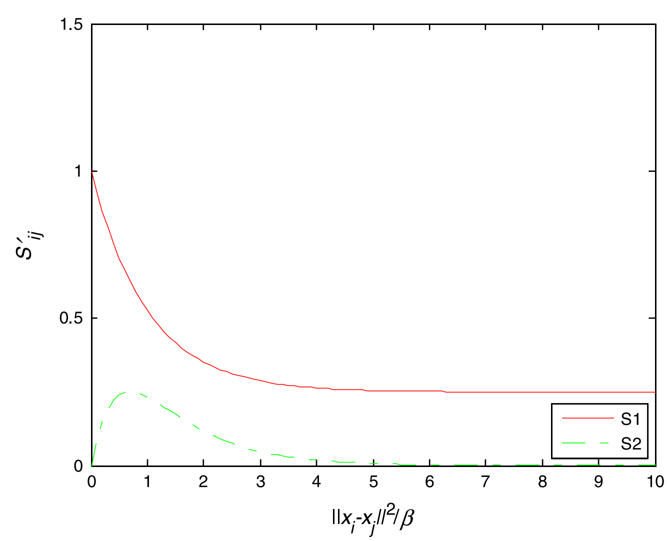

Fig. 1. Plot of S'_{ij} as a function of $||x_i - x_j||^2/\beta$.

* Corresponding author.
E-mail addresses: huangpu3355@163.com (P. Huang), yzcck@126.com, cck.yzu@gmail.com (C. Chen).

http://dx.doi.org/10.1016/j.neucom.2014.03.031
0925-2312/

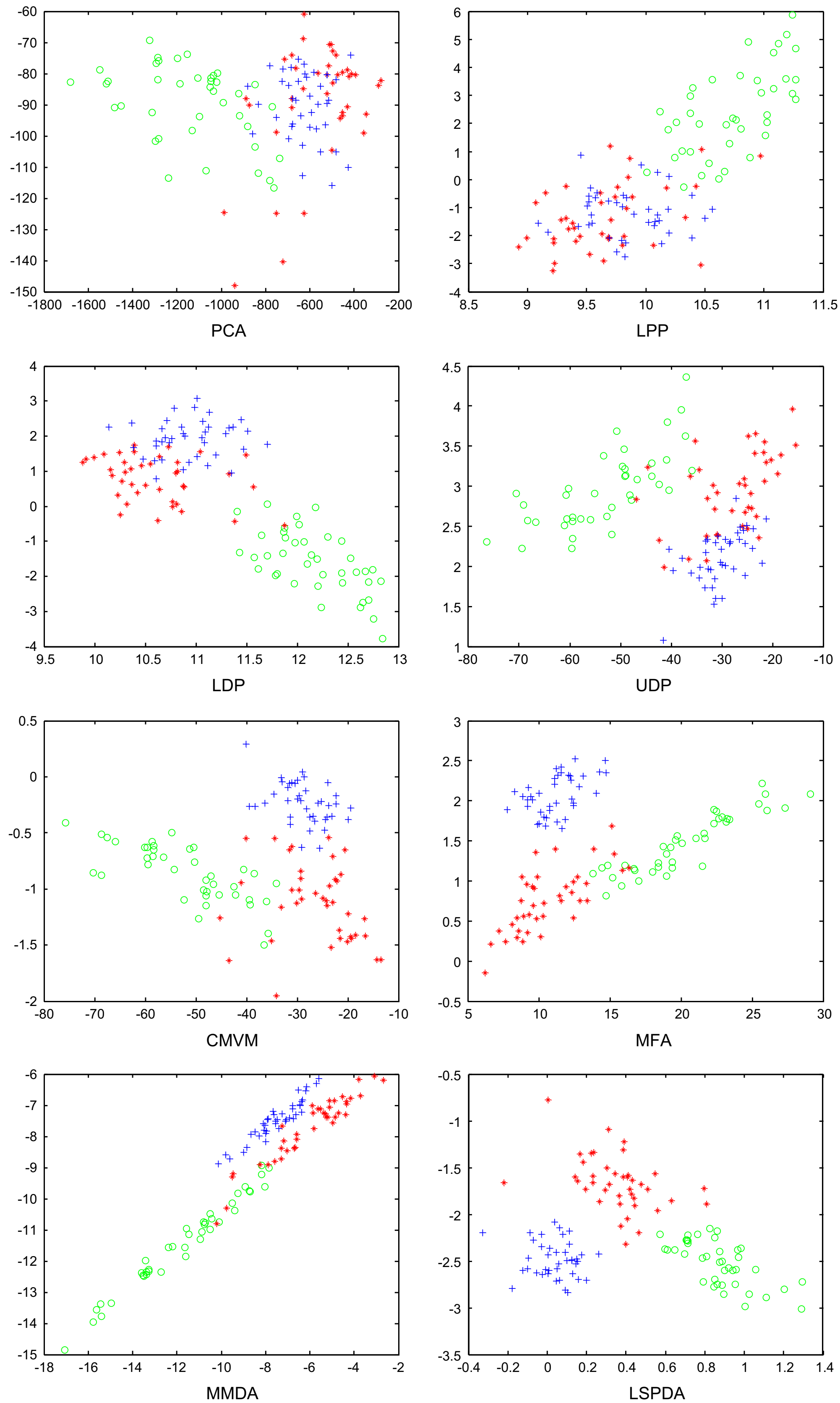

Fig. 2. 2D visualization results on Wine data set by PCA, LPP, LDP, UDP, CMVM, MFA, MMDA and LSPDA, where 'o', '*' and '+' denote three different classes, correspondingly.

conventional LDA which is based on the L_2-norm can degrade with the presence of outliers, Oh et al. proposed an LDA based on the L_p-norm optimization technique (LDA-L_p), which is robust to outliers.

Recent studies [16–18] have shown that large volumes of high-dimensional data possibly reside on a nonlinear sub-manifold. The two classical methods PCA and LDA can see only the global linear