Engineering Applications of Artificial Intelligence 65 (2017) 421–432

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

GMFLLM: A general manifold framework unifying three classic models for dimensionality reduction

Yujin Zhu, Zhe Wang *, Daqi Gao *, Dongdong Li

Department of Computer Science and Engineering, East China University of Science and Technology, Shanghai, 200237, PR China

A R T I C L E I N F O

Keywords:
Pattern recognition
Dimensionality reduction
Locality preserving projection
Linear discriminant analysis
Maximum margin criterion

A B S T R A C T

As one of the most important preprocess in pattern recognition, the dimensionality reduction is widely applied to the real-world tasks. In practice, there exist three corresponding well-known models, including the Locality Preserving Projection (LPP), the Linear Discriminant Analysis (LDA), and the Maximum Margin Criterion (MMC). Even though several previous works have revealed the partial relationship among the three, there are no further researches. In this paper, from the perspective of LPP, the complete connections among the three models are demonstrated, and then a new framework named GMFLLM is proposed to unify them. Further, since it is possible to utilize the proposed framework as an underlying platform to design more dimensionality reduction variants of LPP, fourteen new variants developed from GMFLLM are approached and investigated in the experiment. Moreover, the best of them, named as the Between-class concerned DLPP/MMC (BDLPP/MMC), is selected to compare with the other seven existing state-of-the-art methods on six image datasets. Results validate the effectiveness of BDLPP/MMC so as to show the generalization of GMFLLM.

1. Introduction

Comprehensive applications of manifold learning in pattern recognition have been witnessed in the last decade (Chang et al., 2003; Lee et al., 2003). Developed from the Laplacian Eigenmaps (Belkin and Niyogi, 2001), the Locality Preserving Projection (LPP) is widely used in pattern recognition preprocessing tasks, such as dimensionality reduction (Duda et al., 2012) and graph embedding (Yan et al., 2005, 2007). Differently from the early non-linear techniques, such as the Locally Linear Embedding (Roweis and Saul, 2000) and Isomap (Tenenbaum et al., 2000) that confront the challenge to evaluate the maps on test datasets He et al. (2005), LPP could be unsupervised and pays more attention to preserving the relationship between every two samples. From the perspective of LPP, if two samples locate close to each other in the original space, they should still be close to each other after the projection. As the result, the solution from LPP is concise but powerful. However, two shortcomings remain. Firstly, LPP is sensitive to the neighborhood size. Secondly, LPP suffers from the small sample size problem, which means that the dimensionality of samples is larger than the number of samples so that the data matrix turns to be singular. Since inversing the data matrix is a necessary operation for LPP, overcoming the small sample size problem becomes imperative.

Existing improvements to address the mentioned issues for LPP are categorized into four groups in this paper. The first group contains preprocessings and extensions for LPP, the second one is about the internal improvements, the third one is the external improvements, while the last one is to combine the internal and the external improvements together. Brief retrospect to each group is given in turn below.

(1) For the first group, there are improvements focusing on generating more suitable samples for the subsequent manipulation. For instance, in both of the Diagonal and Secondary Diagonal LPP (DiaLPP & SDiaLPP) (Veerabhadrappa and Rangarajan, 2010) and their extensions, the Diagonal and Weighted Two-Dimensional Discriminant LPP (DiaDLPP & W2D-DLPP) (Lu and Tan, 2011), the original input images are transformed to diagonal ones. This transformation essentially finds a novel way to assign the appropriate weight to each pixel of the original image. Additionally, the other improvements belonging to the group aim to introduce the conventional LPP into new application scenarios. As a trial of bringing LPP into the sub-space learning, the original face images are partitioned to sub-patterns and brought into the Adaptively-weighted Sub-pattern LPP (Aw-SpLPP) (Wang et al., 2010), in order to be separately extracted (Wang et al., 2010).

(2) The second group makes changes under the original model of LPP itself and a majority of the existing improvements in LPP belong to

* Corresponding authors.
E-mail addresses: wangzhe@ecust.edu.cn (Z. Wang), gaodaqi@ecust.edu.cn (D. Gao).

http://dx.doi.org/10.1016/j.engappai.2017.06.028
Received 17 November 2016; Received in revised form 19 May 2017; Accepted 29 June 2017
0952-1976/

this group. They could be further divided into four parts. The methods from the first part improve LPP in the affinity matrix level, i.e., they concern about the distance measurements or the connection constraints between samples. The representative models are listed as follows: For the Bilateral LPP (BLPP) (Li et al., 2015), a filtering term is added behind the Euclidean distance function of the model to balance the weight of each edge of samples; For the Enhanced LPP (ELPP) (Yu et al., 2011), one robust path is introduced into the calculation of the affinity matrix, in order to overcome the sensitivity of LPP to noise and outlier; For the Locally Discriminating Projection (LDP) (Zhao et al., 2006), the supervised way is adopted to build edges of samples. Furthermore, methods from both the second and the third part aim to modify LPP in the model level, but in different ways. One way is to kernelize LPP, containing methods such as the Supervised Kernel LPP (SKLPP) (Cheng et al., 2005) and the Kernel LPP (KLPP) (Feng et al., 2006); the other way is to matrixize LPP, including the Two-Dimensional LPP (2D-LPP) (Chen et al., 2007), the Two-Dimensional Discriminant LPP (2D-DLPP) (Zhi and Ruan, 2008), the Two-Dimensional Discriminant Supervised LPP (2DDSLPP) (Xu et al., 2009), the Sparse Two-Dimensional Discriminant LDP (S2DLDP) (Lai et al., 2011), and the Two-Dimensional Regularized LPP (2DRLPP) (Zhou et al., 2015). Moreover, the methods from the fourth part utilize special optimization ways to overcome the small sample size problem and thus are in the algorithm level. The representative models include the Optimal LPP (OLPP) that transforms the singular eigen-system computation to eigenvalue decomposition problems without losing any discriminative information (Chen et al., 2011), the Exponential LPP (ExLPP) that utilizes the matrix exponential into Laplacian matrix to keep it nonsingular (Wang et al., 2011), and both of the Expression-Specific LPP (ES-LPP) and the Class-Regularized LPP (CR-LPP) that consider priori global and local information to avoid the singularity (Chao et al., 2015).

(3) The third group seeks the cooperation between LPP and the other feature extraction models (He et al., 2005; Lu et al., 2010). Available models include the Linear Discriminant Analysis (LDA) (Belhumeur et al., 1997) and the Maximum Margin Criterion (MMC).[1] According to the degree of the cooperation, these improvements could be further distinguished as two parts. The methods in the first part replace terms in LPP without changing its criterion. For instance, in the method called Discriminant LPP (DLPP) (Yu et al., 2006), the Laplacian term $\boldsymbol{XLX}^T$ of LPP is preserved, while the original diagonal matrix $\boldsymbol{D}$ is replaced with another Laplacian matrix deduced from centroids of classes. Moreover, in the Fisher LPP (FLPP) (Laadjel et al., 2015), a newly-defined matrix $\boldsymbol{L}_b$ is utilized to build the neighbor-graph for samples from different classes . On the other hand, the methods in the second part select the maximum margin based criterion (Li et al., 2006) rather than the generalized rayleigh quotient (Bathe and Wilson, 1976) as the criterion of LPP. One typical improvement is to connect terms generated by DLPP through the maximum margin based criterion, so as to construct the DLPP based on the Maximum Margin Criterion (DLPP/MMC) (Lu et al., 2010). Furthermore, one of the three novel models proposed in the literature (Xu et al., 2010), which is abbreviated as nLPP3 in this paper for convenience, considers to take a similar criterion as MMC to build the framework, in order to avoid the singularity. Practices validate the effectiveness of methods from both parts.

(4) Methods in the fourth group introduce internal improvements into external ones. For instance, according to the Null-space DLPP (NDLPP) (Yang et al., 2008), both the geometrical and discriminant structures of the data manifold are encoded to make the eigenvalue problem solved in null space. Furthermore, in the Regularized Locality Preserving Discriminant Analysis (RLPDA) (Gu et al., 2011), the eigenspace of the locality preserving within-class scatter matrix is first decomposed into three subspaces, and then the three subspaces are regularized differently according to their predicted eigenvalues.

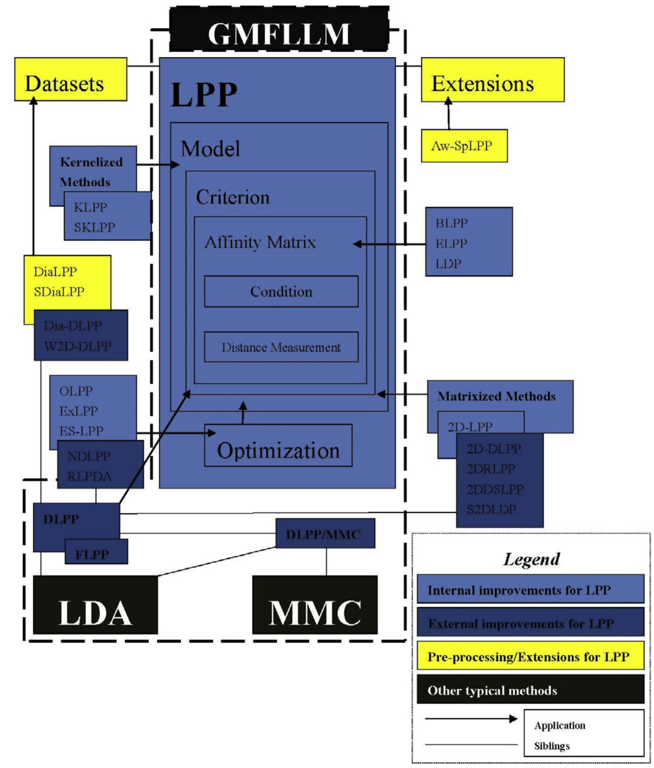

Fig. 1. Relationship graph of LPP-related methods. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 1 is designed to present the relationship among the mentioned LPP-related improvements. In the legend of the figure, "Siblings" means that the methods connected by the line share relationship with each other, while "Application" means that the method could be applied to its corresponding aspect through an arrow. To be clear, different groups of methods are in various colors.

From both the review and the figure, four main points should be seen: (1) The external improvements seem making more essential changes than the internal ones, because the former mainly focuses on the criterion or the term of LPP. (2) The internal improvements seem convenient to be introduced into the external ones. Thus, when designing a well-functioned model for LPP, it is advised to find a suitable criterion first, and then consider the construction of the affinity matrix. (3) As the typical methods for the dimensionality reduction tasks, LDA and MMC are suitable to collaborate with LPP to generate external improvements. (4) LDA and MMC adopt different criterions for their optimization and both of them seem sharing relationship with LPP.

To our best knowledge, there are still not enough researches discussing either the criterion selection of LPP or the complete relationship among LPP, LDA, and MMC (Yan et al., 2005, 2007). Furthermore, there are not any researches proposing a general framework to unify the three main dimensionality reduction models. Since the external improvements seem promising in practice, it becomes necessary to further investigate the condition that combines LPP, LDA and MMC together.

Motivated by the deficiency of the related researches, the complete discussion on the relationship among LPP, LDA, and MMC, will be given in this paper. Thereafter, a general manifold framework abbreviated as GMFLLM is proposed to unify all of the three models. In Fig. 1, GMFLLM is depicted and enclosed by the dashed line. GMFLLM could be distinguished from the existing methods by not only revealing the relationship among some well-known LPP-related models, but also offering chance to develop new variants of LPP for dimensionality

[1] The abbreviation "MMC" is used for both the model and the criterion of the model in Li et al. (2006). To avoid ambiguity, the abbreviation is only used for the model, while the full name "the maximum margin based criterion" is used for the corresponding criterion in this paper.